Analysis of complex tannins by multidimensional techniques

by
Pieter Venter

Thesis presented in partial fulfilment of the requirements for the degree of
Doctor of Philosophy (PhD)

at
Stellenbosch University
Department of Chemistry and Polymer Science

Supervisor: Prof. André de Villiers
Co-supervisor: Prof. Harald Pasch

April 2019

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2019

Abstract

Hydrolysable tannins are plant derived (poly)phenolics which for centuries have been used in tanning, the transformation of animal hides into leather. In addition to their industrial applications, epidemiological evidence suggests that the consumption of hydrolysable tannins can be linked to disease prevention and treatment, although their mode of action at a molecular level is still unknown. In order to investigate their biological and chemical properties, the determination of the molecular composition of hydrolysable tannins in plant material is a critical step. Liquid chromatography hyphenated to mass spectrometry (LC-MS) is the preferred analytical approach to obtain this information, although conventional methods often provide insufficient performance for highly complex mixtures of hydrolysable tannins.

The goal of the work reported in this thesis was, therefore, to develop improved multidimensional methods for the analysis of hydrolysable tannins. In the first part of the study, a methodology for the comprehensive analysis of ellagitannins and gallotannins in chestnut and tara using reversed phase liquid chromatography (RP-LC) and hydrophilic interaction chromatography (HILIC) combined with ultraviolet (UV), ion mobility (IM) spectrometry and high-resolution mass spectrometry (HR-MS) detection is reported. Good chromatographic performance was achieved using both HILIC and RP-LC for chestnut tannins, with the complementary separation mechanisms proving useful for particular compound classes also pointing to the potential of the combination of HILIC and RP-LC in a comprehensive two-dimensional separation. In contrast, RP-LC provided much better separation of tara gallotannins than HILIC. Incorporation of IM into the LC-MS workflows offered several advantages for such non-targeted screening analyses, including improved mass spectral quality by filtering MS data according to IM arrival time, the availability of an additional identification criterion in the form of collisional cross section (CCS) values obtained from arrival times, and an improved MS duty cycle and, therefore improved sensitivity. Furthermore, additional isomeric species were resolved by IM and, in combination with UV spectral data, IM provided a simple methodology to differentiate between positional isomers.

In the second part of the study, an on-line comprehensive two-dimensional HILIC×RP-LC method was developed for the analysis of hydrolysable tannins by kinetic

optimisation of the relevant experimental parameters. The use of a weak make-up flow to dilute the HILIC effluent and avoid second dimension injection band broadening resulted in excellent performance, with practical peak capacities in excess of 1000 determined for both samples. Despite on-column dilution and the short second dimension analysis times ultimately limiting the number of compounds detected in chestnut and tara, the group-type separations obtained show promise for the fast fingerprint analysis of hydrolysable tannins.

Building on previous findings, the final part of the work evaluated the potential of IM spectrometry as a third dimension in a comprehensive three-dimensional HILIC×RP-LC×IM configuration. Apart from the inherent advantages of IM alluded to above, the additional separation step offered by IM contributed to an increase in practical peak capacity by a factor of 13 when integrated into the LC×LC workflow. Certain instrumental constraints which limit the ultimate performance of LC×LC×IM-HR-MS are highlighted. Despite these, this approach shows significant promise for the high resolution separation of highly complex mixtures.

Opsomming

Hidroliseerbare tanniene is plant gebaseerde (poli)fenole wat al eeue lank gebruik word in looiery, die omskakeling van diere velle in leer. Bykomend tot hul industriële toepassings dui epidemiologiese bewyse daarop dat die inname van hidroliseerbare tanniene gekoppel kan word aan die voorkoming en behandeling van siektes, alhoewel hul werkswyse op molekulêre vlak nog onbekend is. Ten einde hul biologiese en chemiese eienskappe te ondersoek, is die bepaling van die molekulêre samestelling van hidroliseerbare tanniene in plantmateriaal 'n kritieke stap. Vloeistofchromatografie gekoppel aan massa spektrometrie (LC-MS) is die analitiese benadering wat meestal gebruik word om hierdie inligting te verkry, hoewel konvensionele metodes dikwels onvoldoende resultate vir hoogs komplekse mengsels van hidroliseerbare tanniene lewer.

Die doel van die werk wat in hierdie proefskrif gerapporteer is, was dus om verbeterde multidimensionele metodes vir die analise van hidroliseerbare tanniene te ontwikkel. In die eerste gedeelte van die studie is 'n metodologie vir die omvattende analise van ellagitanniene en gallotanniene in chestnut en tara deur die gebruik van omgekeerde-fase vloeistofchromatografie (RP-LC) en hidrofiliese interaksie chromatografie (HILIC) gekombineer met ultraviolet (UV), ioon mobiliteit (IM) spektrometrie en hoë-resolusie massa spektrometrie (HR-MS) deteksie gerapporteer. Goeie chromatografiese skeiding is behaal met die gebruik van beide HILIC en RP-LC vir chestnut tanniene, met die komplementêre skeidingsmeganismes wat nuttig bewys is vir spesifieke komponent klasse, en dus ook dui op die potensiaal van die kombinasie van HILIC en RP-LC in 'n omvattende twee-dimensionele skeiding. In teenstelling hiermee lewer RP-LC ‘n baie beter skeiding van tara gallotanniene as HILIC. Die inkorporering van IM in die LC-MS werksvloei het verskeie voordele vir sulke nie-geteikende analise. Dit sluit in verbeterde massaspektrum kwaliteit deur MS-data volgens IM-aankomstyd te filtreer, die beskikbaarheid van 'n bykomende identifikasie parameter in die vorm van botsing-dwarssnit (CCS) waardes verkry vanaf aankoms tye, en 'n meer doeltreffende massa spektrometer wat lei tot beter sensitiwiteit. Verder, isomeriese spesies wat deur IM geskei kan word, in kombinasie met UV spektrale data, verskaf 'n eenvoudige metode om tussen posisionele isomere te onderskei.

In die tweede deel van die studie is 'n aanlyn-omvattende twee-dimensionele HILIC×RP-LC metode vir die analise van hidroliseerbare tanniene deur kinetiese optimalisering van die relevante eksperimentele parameters ontwikkel. Verdunning van die eerste dimensie (^{1}D) eluent met swak tweede dimensie (^{2}D) mobiele fase om inspuitings-bandverbreding in die ^{2}D te beperk het uitstekend skeiding gelewer, met praktiese piek kapasiteite van meer as 1000 wat vir beide monsters bepaal is. Ten spyte van in-kolom verdunning en kort tweede dimensie analise tye wat die aantal verbindings wat in chestnut en tara opgespoor word beperk, toon die groep-tipe skeidings potensiaal vir die vinnige vingerafdruk analise van hidroliseerbare tanniene.

Gebaseer op vorige bevindinge, het die finale deel van die werk die potensiaal van IM-spektrometrie as 'n derde dimensie in 'n omvattende drie-dimensionele HILIC×RP-LC×IM opset geëvalueer. Afgesien van die inherente voordele van IM wat hierbo genoem word, het die bykomende skeidingsstap wat deur IM gelewer word die praktiese piek kapasiteit met ‘n faktor van 13 verhoog wanneer dit in die LC×LC werksvloei geïntegreer is. Verskeie instrumentele limitasies wat die uiteindelike prestasie van LC×LC×IM-HR-MS beperk, word uitgelig. Ten spyte hiervan toon hierdie benadering betekenisvolle potensiaal vir die hoë resolusie skeiding van hoogs komplekse mengsels.

Acknowledgements

Foremost, I would like to show my appreciation for Prof. André de Villiers for giving me the opportunity to obtain my PhD. André, your day-to-day behaviour can be described as someone who leads by example. It was through your example of hard work and humility that I have progressed not only as a scientist, but also as a person.

In the same breath, my appreciation goes to Prof. Harald Pasch, who, through his financial support, made this journey possible in the first place.

A special thanks to Magriet Muller whose natural ability to grasp this subject has added considerable knowledge to my project. On the same note I would like to thank Jochen Vestner for his visual contribution to my thesis.

Thank you, Dr. Marietjie Stander, for allowing us to have carte blanche in the MS lab.

I would also like to show my gratitude towards my fellow colleagues, Prof Burger, Andreas, Aron, Sithandile, Elizabeth, Tlou, Magriet and Ester for creating a pleasant working environment and a helping hand, as well as for Shafiek for doing the administrative work.

I gratefully appreciate the financial support from the National Research Foundation and Sasol.

Vir my familie is ek baie dankbaar vir julle volgehoue ondersteuning en geloof in my.

List of Abbreviations

APCI: atmospheric pressure chemical ionization

β: undersampling factor

CCS/Ω: collisional cross section

$^{DT}CCS_{N2}$: collisional cross section determined on a drift tube ion mobility spectrometer with nitrogen as collision gas

$^{TW}CCS_{N2}$: collisional cross section determined on a travelling wave ion mobility spectrometer with nitrogen as collision gas

CID: collision induced dissociation

CIU: collision-induced unfolding

CPC: centrifugal partitioning chromatography

CQA: caffeoyl quinic acid

^{1}D: first dimension

^{2}D: second dimension

DMA: differential mobility analyser

DTIMS: drift time ion mobility spectrometry

ESI: electrospray ionization

FAIMS: field asymmetric waveform ion mobility spectrometry

f_c: fractional surface coverage

GPC: gel permeation chromatography

GQA: galloyl quinic acids

HHDP: hexahydroxydiphenoyl

HILIC: hydrophilic interaction chromatography

HILIC×RP-LC: two-dimensional hydrophilic interaction chromatography × reversed phase liquid chromatography

HILIC×IM: hydrophilic interaction chromatography × ion mobility spectrometry

HR-MS: high resolution mass spectrometry

HPLC: high performance liquid chromatography

IM: ion mobility spectrometry

IM-MS: ion mobility-mass spectrometry

LC×LC: comprehensive two-dimensional liquid chromatography

MALDI: matrix assisted laser desorption/ionization

MD: multidimensional

MRM: multiple reaction monitoring

MS: mass spectrometry

MS^E: high energy mass spectrum

ms: milliseconds

m/z: mass-to-charge ratio

n_c: peak capacity

^1n_c: first dimensional peak capacity

^2n_c: second dimensional peak capacity

^3n_c: third dimensional peal capacity

n'_c,: corrected peak capacity

NHTP: nonahydroxyterphenoyl

NMR: nuclear magnetic resonance

NP-LC: normal phase liquid chromatography

ppm: parts per million

Q-TOF: quadrupole time-of-flight mass spectrometry

RP-LC: reversed phase liquid chromatography

RP-LC×IM: reversed phase liquid chromatography × ion mobility spectrometry

RSD: relative standard deviation

SEC: size exclusion chromatography

SOD1: superoxide dismutase

t_d: recorded arrival (drift) time

$t_{d,cycle}$: ion mobility cycle time

t_g: gradient time

TIMS: trapped ion mobility spectrometry

TOF: time-of-flight mass spectrometry

t_R: retention time

TWIMS: travelling wave ion mobility spectrometry

UHPLC: ultra-high pressure liquid chromatography

UV: ultraviolet spectroscopy

Table of contents

Chapter 1

General introduction and objectives

1.1 General introduction

Hydrolysable tannins are a sub-class of the (poly)phenolic family that are ubiquitous secondary metabolites in the plant kingdom, and are important constituents of the human diet [1]. Hydrolysable tannins can be broadly grouped into two sub-classes, namely gallotannins and ellagitannins, with each class comprising a large number of compounds ranging in molecular weight between ~300 and more than 5000 Dalton. More than 1000 hydrolysable tannins have been identified to date [2]. Hydrolysable tannins have for centuries been used in the conversion of hide into leather, a process called tanning, where the phenolic hydrolysable tannins form hydrogen bonds with collagen to improve the properties and stability of the leather [3]. More recent industrial applications of hydrolysable tannins include the production of wood panel adhesives [4–6]. Furthermore, significant interest in hydrolysable tannins in recent years can be ascribed to the reported health benefits associated with their consumption, which is based on epidemiological evidence [7].

To understand the mechanism(s) by which hydrolysable tannins exert their health benefits, or the function they perform in the commercial applications of plant-derived commodities, their quantitative and qualitative analytical characterisation in natural products is essential. Indeed, a range of techniques have been applied for the analysis of hydrolysable tannins. Nuclear magnetic resonance (NMR) spectroscopy following preparative isolation is undoubtedly the method of choice for the structural elucidation of these compounds, and has played a pivotal role in this regard [8,9]. However, for complex mixtures of hydrolysable tannins, which often contain trace level constituents, this approach is not suitable. To obtain detailed information about the hydrolysable tannin composition of such samples, the combination of analytical scale chromatographic separation with ultraviolet (UV) spectroscopy and mass spectrometry (MS) has become an irreplaceable tool. Reversed phase liquid chromatography (RP-LC) in particular is almost exclusively employed for the analysis of hydrolysable tannins in natural products and foodstuffs [2], as well as their metabolites as they pass through the gastrointestinal tract [10–13]. UV (diode array) detection can be used to distinguish between subspecies of hydrolysable tannins [14], as well as between regioisomers [15]. MS has become an indispensible tool in hydrolysable tannin analysis, both for targeted analysis (typically using tandem MS

instruments [16] and for non-targeted analysis. For the latter type of analyses, high resolution mass spectrometry (HR-MS) is especially useful to reveal structural data which can be used for (tentative) identification of known and unknown hydrolysable tannins [15]. However, due to the sheer complexity of hydrolysable tannin mixtures such as encountered in natural product and food related products, contemporary LC-UV-MS methods suffer from incomplete separation and failure to identify trace-level co-eluting compounds.

One way to address separation challenges is to combine different (orthogonal) separation methods. Such multidimensional (MD) separation systems are capable of significantly improving performance regarding the number of peaks resolved [17]. Comprehensive two-dimensional liquid chromatography (LC×LC) in particular has in recent years shown promise for the improved separation of complex mixtures of phenolic compounds [18,19] due to the much higher resolving power the technique offers for conventional separation times. Indeed, while LC×LC has been successfully used for the analysis of condensed tannins, mainly using HILIC and RP-LC [18], it has not to date been applied to hydrolysable tannin analysis.

An alternative approach to obtain improved MS data, which has gained significant traction in recent years, is to use ion mobility (IM) spectrometry. Ion mobility spectrometry is the gas phase separation of ions achieved by the collision with an inert buffer gas while under the influence of an electrical field [20]. Benefits of using ion mobility spectrometry include an increased MS duty cycle for better sensitivity [21], deconvolution of multiple charged species [22], as well as offering different fragmentation options [23].

With the above as background, the aim of this study was to explore the potential of novel hyphenated analytical techniques for the detailed chemical analysis of hydrolysable tannins. In particular, the utility of IM-HR-MS for this application will be evaluated in combination with one- and multidimensional LC separations. Considering their widespread application in phenolic analysis, HILIC and RP-LC will be used as separation methods. While HILIC has not to date been applied for hydrolysable tannin analysis, the complementary separation mechanism of this LC mode is expected to provide complementary information to RP-LC. As application, chestnut and tara

samples will be used, as these contain predominantly ellagitannins and gallotannins, respectively.

1.2 Aims and objectives

This study was aimed at developing novel analytical methodologies for the comprehensive analysis of complex mixtures in commercial tannin samples using multidimensional separation techniques. To achieve this, the following objectives were formulated:

1. To explore the potential of complementary separation modes, in the form of HILIC and RP-LC, hyphenated to UV and IM-HR-MS detection for the detailed analysis of chestnut and tara hydrolysable tannins. The benefits of incorporating IM into the LC-MS workflow will also be evaluated for this application.
2. The second objective was to exploit the benefits of LC×LC separation for hydrolysable tannin analysis by developing an on-line comprehensive two-dimensional HILIC×RP-LC-UV-HR-MS method for the analysis of chestnut and tara tannins.
3. Building on the findings of the precedent work, the final objective was to hyphenate HILIC×RP-LC separation to IM-HR-MS to provide a comprehensive three-dimensional separation system for the detailed analysis of complex phenolic mixtures, including hydrolysable tannins. As part of this work, the advantages and limitations of IM in such an LC×LC×IM-MS system will also be evaluated.

1.3 References

1. Serrano J, Puupponen-Pimiä R, Dauer A, Aura AM, Saura-Calixto F (2009) Tannins: Current knowledge of food sources, intake, bioavailability and biological effects. Mol Nutr Food Res 53:310–329. doi: 10.1002/mnfr.200900039

2. Arapitsas P (2012) Hydrolyzable tannin analysis in food. Food Chem 135:1708–1717. doi: 10.1016/j.foodchem.2012.05.096

3. White T (1957) Tannins-Their occurrence and significance. J Sci Food Agric 8:377–385. doi: 10.1002/jsfa.2740080702

4. Spina S, Zhou X, Segovia C, Pizzi A, Romagnoli M, Giovando S, Pasch H, Rode K, Delmotte L (2013) Phenolic resin adhesives based on chestnut (Castanea sativa) hydrolysable tannins. J Adhes Sci Technol 27:2103–2111. doi: 10.1080/01694243.2012.697673

5. Aouf C, Benyahya S, Esnouf A, Caillol S, Boutevin B, Fulcrand H (2014) Tara tannins as phenolic precursors of thermosetting epoxy resins. Eur Polym J 55:186–198

6. Abdalla S, Pizzi A, Bahabri F, Ganash A (2015) Analysis of Valonia Oak (Quercus aegylops) Acorn Tannin and Wood Adhesives Application. Bioresources 10:7165–7177

7. Crozier A, Jaganath IB, Clifford MN (2009) Dietary phenolics: Chemistry, bioavailability and effects on health. Nat Prod Rep 26:1001–1043. doi: 10.1039/b802662a

8. Glabasnia A, Hofmann T (2006) Sensory-Directed Identification of Taste-Active Ellagitannins in American (Quercus alba L.) and European Oak Wood (Quercus robur L.) and Quantitative Analysis in Bourbon Whiskey and Oak-Mutured Red Wines. J Agric Food Chem 54:3380–3390. doi: 10.1021/jf052617b

9. Suvanto J, Tähtinen P, Valkamaa S, Engström MT, Karonen M, Salminen JP (2018) Variability in Foliar Ellagitannins of Hippophaë rhamnoides L. and Identification of a New Ellagitannin, Hippophaenin C. J Agric Food Chem 66:613–620. doi: 10.1021/acs.jafc.7b04834

10. Piwowarski JP, Stanisławska I, Granica S, Stefanska J, Kiss AK (2017) Phase II conjugates of urolithins isolated from human urine and potential role of β-glucuronidases in their disposition. Drug Metab Dispos 45:657–665. doi: 10.1124/dmd.117.075200

11. Cerdá B, Llorach R, Cerón JJ, Espín JC, Tomás-Barberán FA (2003) Evaluation of the bioavailability and metabolism in the rat of punicalagin , an antioxidant polyphenol from pomegranate juice. Eur J Nutr 42:18–28. doi: 10.1007/s00394-003-0396-4

12. Cerdá B, Espín JC, Parra S, Martínez P, Tomás-Barberán FA (2004) The potent in vitro antioxidant ellagitannins from pomegranate juice are metabolised into bioavailable but poor antioxidant hydroxy-6H-dibenzopyran-6-one derivatives by the colonic microflora of healthy humans. Eur J Nutr 43:205–220. doi: 10.1007/s00394-004-0461-7

13. Cerdá B, Tomás-Barberán FA, Espin JC (2005) Metabolism of Antioxidant and Chemopreventive Ellagitannins from Strawberries, Raspberries, Walnuts, and Oak-Aged Wine in Humans: Identification of Biomarkers and Individual Variability. J Agric Food Chem 53:227–235. doi: 10.1021/jf049144d

14. Moilanen J, Sinkkonen J, Salminen JP (2013) Characterization of bioactive plant ellagitannins by chromatographic, spectroscopic and mass spectrometric methods. Chemoecology 23:165–179. doi: 10.1007/s00049-013-0132-3

15. Arapitsas P, Menichetti S, Vincieri FF, Romani A (2007) Hydrolyzable tannins with the hexahydroxydiphenoyl unit and the m-depsidic link: HPLC-DAD-MS identification and model synthesis. J Agric Food Chem 55:48–55

16. Engström MT, Pälijärvi M, Salminen JP (2015) Rapid fingerprint analysis of plant extracts for ellagitannins, gallic acid, and quinic acid derivatives and quercetin-, kaempferol- and myricetin-based flavonol glycosides by UPLC-QqQ-MS/MS. J Agric Food Chem 63:4068–4079. doi: 10.1021/acs.jafc.5b00595

17. Giddings JC (1984) Two-dimensional separations: concept and promise. Anal Chem 56:1258A–1270A. doi: 10.1021/ac00276a003

18. Villiers A de, Venter P, Pasch H (2016) Recent advances and trends in the liquid-chromatography – mass spectrometry analysis of flavonoids. J Chromatogr A

1430:16–78. doi: 10.1016/j.chroma.2015.11.077

19. Cacciola F, Farnetti S, Dugo P, Marriott PJ, Mondello L (2016) Comprehensive Two-dimensional Liquid Chromatography for Polyphenol Analysis in Foodstuffs. J Sep Sci In press:. doi: 10.1002/jssc.201600704.

20. Clemmer DE, Jarrold MF (1997) Ion mobility measurement and their applications to cluster and biomolecules. J Mass Spectrom 32:577–592

21. Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH (2004) Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. Rapid Commun Mass Spectrom 18:2401–2414. doi: 10.1002/rcm.1641

22. Pringle SD, Giles K, Wildgoose JL, Williams JP, Slade SE, Thalassinos K, Bateman RH, Bowers MT, Scrivens JH (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. Int J Mass Spectrom 261:1–12. doi: 10.1016/j.ijms.2006.07.021

23. Sun J, Baker A, Chen P (2011) Profiling the indole alkaloids in yohimbe bark with ultra-performance liquid chromatography coupled with ion mobility quadrupole time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 25:2591–2602. doi: 10.1002/rcm.5158

Chapter 2

Literature review

2.1 Hydrolysable tannins: General chemistry and importance

Hydrolysable tannins are secondary metabolites derived by the metabolism of the gallic acid (itself derived from the shikimate pathway [1]), which occur widespread throughout the plant kingdom [2]. These compounds are, therefore, a sub-class of polyphenols, recently defined by Quideau and co-workers [1] as "*plant secondary metabolites derived exclusively from the shikimate-derived phenylpropanoid and/or the polyketide pathway(s), featuring more than one phenolic ring and being devoid of any nitrogen-based functional group in their most basic structural expression"*.

The term “hydrolysable tannin” originates from the hydrolytic cleavage of the polyol core and the acid moiety of tannins in acidic [3, 4] or basic solution [5]. Hydrolysable tannins were defined by White, Bate-Smith, Swain and Haslam (WBSSH) as polyphenols which are water soluble, in the mass range of 500 to 3000-4000 Da, containing twelve to sixteen phenolic groups located on five to seven aromatic rings per 1000 relative molecular mass [6] and possess the ability to convert animal skins into leather, in a process called tanning. Hydrolysable tannins can broadly be grouped into gallotannins and ellagitannins. **Figure 2.1** illustrates the structural differences between these classes.

The ellagitannins vescalagin and tellimagrandin I both contain a hexahydroxydiphenoyl (HHDP) group that is esterified to a glucose (polyol) core (**Figure 2.1A** and **B**) [7]. A nonahydroxyterphenoyl (NHTP) unit can also be esterified to the polyol core, as is the case for vescalagin (**Figure 2.1A**) [8]. The glucose core in ellagitannins can either be in the open or closed configuration, illustrated by vescalagin (**Figure 2.1A**) and tellimagrandin I (**Figure 2.1B**), respectively [9].

In contrast, gallotannins contain galloyl groups esterified to either a closed glucose or a quinic acid polyol core. Examples of gallotannins are trigalloyl glucose (**Figure 2.1C**) and 1-O-digalloyl-3,4,5-tri-O-galloylquinic acid (**Figure 2.1D**), respectively [10]. The galloyl moiety can furthermore be esterified to another gallic acid group through a depsidic bond [6], as illustrated for the 1-O-digalloyl group in **Figure 2.1D**.

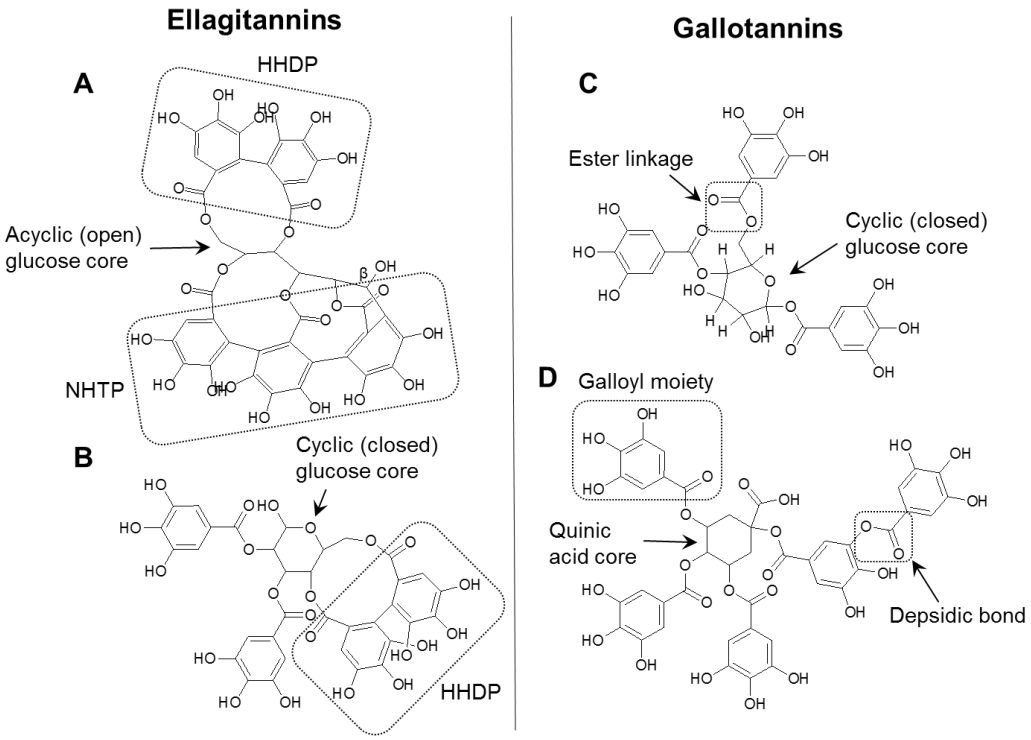


Figure 2.1: Structural differences between ellagitannins and gallotannins, which together comprise the hydrolysable tannins. Ellagitannins are represented by vescalagin (**A**, open glucose core) and tellimagrandin I (**B**, closed glucose core) and gallotannins by trigalloyl glucose (**C**) and 1-O-digalloyl-3,4,5-tri-O-galloylquinic acid (**D**). Abbreviations: HHDP: hexahydroxydiphenoyl, NHTP: nonahydroxyterphenoyl.

Hydrolysable tannins are of interest due to their industrial applications, as well as the biological roles they play in plants and living organisms. Historically, their primary application has been in the conversion of animal hides into leather; plant-derived hydrolysable tannins have been used in tanning for more than a thousand years [1]. Furthermore, hydrolysable tannins have been used as renewable alternatives to petroleum-derived phenols in the manufacturing of resin adhesives [11–14] and rigid foams [15]. For example, chestnut hydrolysable tannins can replace a large fraction (60%) of petroleum-derived phenol in the production of adhesives while still meeting the specifications required for particular applications [11]. Further industrial applications of hydrolysable tannins include the use of chestnut wood barrels (cooperage) for the production of wine and spirits [16], and as rust converter [17].

Known biological roles of hydrolysable tannins include their defensive function against herbivores in plants as well as their potential health benefits to humans. Hydrolysable tannins are thought to act as defensive compounds in plants, either through their pro-oxidative properties [18] or their ability to precipitate proteins [19, 20]. The pro-oxidant hypothesis was based on a deoxyribose assay conducted by Moilanen *et al*. [21], which showed that hydrolysable tannins undergo oxidative reactions resulting in pro-oxidants (semiquinones and quinones) under conditions (pH 10) similar to those encountered in the digestive midguts of insect herbivores. However, Moilanen *et al*. [21] and Gyamfi and Aniya [22] also reported the antioxidant ability of ellagitannins to chelate Fe^{2+} and Fe^{3+}, thus preventing them from participating in radical formation.

Much of the contemporary interest in hydrolysable tannins by the scientific community, as well as in the public domain, is due to their abundant occurrence in many dietary foodstuffs whose consumption has been associated with potential health benefits [2]. Beneficial properties ascribed to polyphenols and hydrolysable tannins include their supposed capability to act as biological (*in vivo*) antioxidants to prevent oxidative stress resulting in age related diseases, although this theory has led to much controversy [1]. There is evidence that suggests that once hydrolysable tannins are metabolised in the gut, they lose their powerful *in vitro* antioxidant activity [23–25]. Because of this, it has been proposed that ellagic acid, the hydrolysis product of hydrolysable tannins, and/or its bioavailable metabolites called urolithins, could be the agents responsible for the health promoting effects of hydrolysable tannins tested *in vitro* [26–28]. However, to date there is no conclusive proof of the actual mechanism(s) by which urolithins are responsible for the health benefits to humans ascribed to hydrolysable tannin consumption [29]. Apart from their redox properties, abundant literature reports have shown that polyphenols in general, and hydrolysable tannins in particular, have the ability to target specific proteins (enzymes) and peptides, which can either affect signal transduction pathways or inhibit various enzymatic reactions [1]. This mode of action, which is only proven *in vitro*, has shown results in the fight against chronic diseases, such as cancer [30, 31] and neurodegenerative diseases [32]. Hydrolysable tannins have been well studied, as is the case for condensed tannins, for their *in-vitro* inhibitory activity against a number of infectious viruses [33–36]. Among 13 tannin samples containing both hydrolysable and condensed tannins, chestnut showed the lowest minimum inhibitory concentration against HIV

cytopathogenicity. Furthermore, although a combination of chestnut and tara tannins showed similar inhibitory activity against the influenza virus as condensed tannin samples, a lower minimum cytotoxic concentration was measured for the former sample [36]. Finally, hydrolysable tannins have also shown to be potential inhibitors of α-glucosidase in the treatment of diabetes [37].

In order to investigate the purported health properties from *in vitro/ex vivo* studies, it is important to assess the bioavailability and metabolism of hydrolysable tannins as they pass through the gastrointestinal tract. The first study to determine the bioavailability of hydrolysable tannins was conducted on rats, and confirmed the presence of trace amounts of the hydrolysable tannin punicalagin in plasma after consumption of pomegranate [38]. This was, however, in contrast to studies conducted on human subjects, which showed no sign of native hydrolysable tannins in plasma or urine after the consumption of either pomegranates or raspberries [24, 39, 40]. Nevertheless, several studies have concluded that ellagitannins are metabolised by colon microbiota to bioavailable urolithins [40–42]. Furthermore, by analysing ileal fluid from subjects with an ileostomy (i.e. without passing through the colon), recoveries of 23% and 241% of sanguiin H-6 and ellagic acid, respectively, were measured after the consumption of raspberries [40]. This study, therefore, revealed extensive hydrolysis of ellagitannins in the stomach and small intestine. To overcome the low bioavailability of phenolics, a series of drug delivery systems has been investigated [43].

2.2 Analysis of hydrolysable tannins

In order to study the possible commercial use and health benefits of hydrolysable tannins, it is essential to determine their chemical composition in biological and natural samples. To achieve this, a typical workflow involves pre-extraction with a non-polar solvent to remove lipids, followed by an aqueous-organic extraction of tannins and several chromatographic isolation steps. Characterisation of the final isolate is usually achieved using nuclear magnetic resonance (NMR), ultraviolet (UV) and mass spectrometry (MS).

Aqueous-organic solvent mixtures used for the extraction of hydrolysable tannins include, in addition to water, methanol, ethanol or acetone, while ethyl acetate has

also found application for the extraction of tannins [44]. For gallotannins in particular, ethyl acetate is the preferred solvent for the Soxhlet extraction of high molar mass components, whereas water is preferred for low molar mass components [45, 46]. Ethanol is a suitable intermediate solvent for the extraction of both low and high molar mass gallotannins, although providing lower yields compared to ethyl acetate and water. For ellagitannins, on the other hand, the most effective extraction solvent is methanol, with water, acetone and ethanol showing equal but lower extraction efficiencies compared to methanol [47, 48]. Cam *et al*. [48] demonstrated that pressurised water (102.1 atm) can be as effective as methanol for the extraction of ellagitannins from pomegranate peels.

Seminal work on the chromatographic analysis of hydrolysable tannins by Okuda and co-workers [49–52] revealed a size-based separation of oligomeric tannins by normal phase liquid chromatography (NP-LC), compared to a more complicated isomeric separation obtained by reversed phase liquid chromatography (RP-LC). A method developed by Scalbert and co-workers [53], which included size exclusion chromatography (SEC) separation followed by preparative RP-LC, was used for characterisation of ellagitannins in chestnut and oak using NMR spectroscopy. Compounds characterised using this workflow were the abundant hexahydroxydiphenoyl (HHDP) species such as the hydrosoluble anomeric ellagitannins vescalagin and castalagin, and their dimeric derivatives (roburins A-E). Recent analysis of oak-aged wine [54] revealed partial co-elution of the dimeric species roburins A, B and C on a fused-core RP-LC column following a preparative SEC separation similar to the method employed by Scalbert and co-workers. The same chromatographic workflow comprising SEC fractionation followed by preparative RP-LC has been used to isolate ellagitannins from several sample matrixes, including whiskey and wine after cooperage in oak wood [55], sea-buckthorn [56], purple loosestrife [57], as well as for the isolation of gallotannins from mango kernels [58]. Incidentally, the whiskey constituents responsible for imparting an astringent oral sensation were identified as the monomeric ellagitannins vescalagin and castalagin and their dimeric counterparts, roburins A-E using MS and NMR [55]. Kool *et al*. [59] isolated four ellagitannins from boysenberry, namely Sanguiin H-2, Sanguiin H-6, Sanguiin H-10 and galloyl Sanguiin H-6, by employing ion exchange chromatography followed by two preparative C18 separations. The isolated compounds were identified

using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) and NMR spectroscopy.

By far the most important analytical technique for structural elucidation of hydrolysable tannins is nuclear magnetic resonance (NMR) spectroscopy. For example, the general size exclusion chromatography (SEC) and RP-LC extraction protocol alluded to above has been employed in combination with NMR for the unambiguous identification of several hydrolysable tannins in different plant species [53, 55–57, 59]. This and similar approaches, however, require elaborate extraction and clean-up steps to provide sufficient amounts (>5 mg) of suitably purified compounds. Analytical-scale separations are more suitable for screening of complex mixtures of hydrolysable tannins without the need to purify individual compounds. Since this work focuses on analytical-scale characterisation of hydrolysable tannins, the following discussion will be limited to the relevant techniques used, which include LC hyphenated to UV and MS detection.

Gallotannins and ellagitannins can be differentiated using UV spectroscopy based on absorption bands at 270-280 nm and 200-220 nm, respectively (**Figure 2.2**) [44, 60–62]. Also, mixed hydrolysable tannin species containing both galloyl and HHDP units in different ratios can be differentiated based on the relative intensity of the absorption band between 273-278 nm [61, 62]. Ellagic acid, the hydrolysis product of HHDP-containing ellagitannins, has a unique UV spectrum, with absorption bands at 254 nm and 360 nm (**Figure 2.2-G**) [60, 63]. Furthermore, gallotannins containing *m*-depsidic bonds can be identified by a shoulder at 300 nm in their UV spectra, which becomes more pronounced with increasing number of depsidic bonds [64].

Mass spectrometry is an essential tool in the analysis of hydrolysable tannins, both as a standalone technique and in combination with chromatographic separation. A wide range of mass spectrometers have been used for this purpose, including single- and multi-stage systems and both nominal mass and high-resolution instruments. Single quadrupole, triple quadrupole, ion trap, time-of-flight (TOF), quadrupole time-of-flight (Q-TOF) and Orbitrap mass spectrometers have been used in the analysis of hydrolysable tannins [44, 65]. Negative electrospray ionisation is the primary ionisation technique used for mass spectrometric analysis of hydrolysable tannins [44], whereas positive electrospray ionisation has been used to a lesser extent [66–69]. There has

been no comparative study on which ionisation mode provides optimal detection sensitivity for hydrolysable tannins. However, it has been demonstrated that there is no significant difference in ionisation potential between positive and negative electrospray ionisation for condensed tannins, but that operating in negative ion mode decreases the background noise and thereby increases the sensitivity [70, 71].

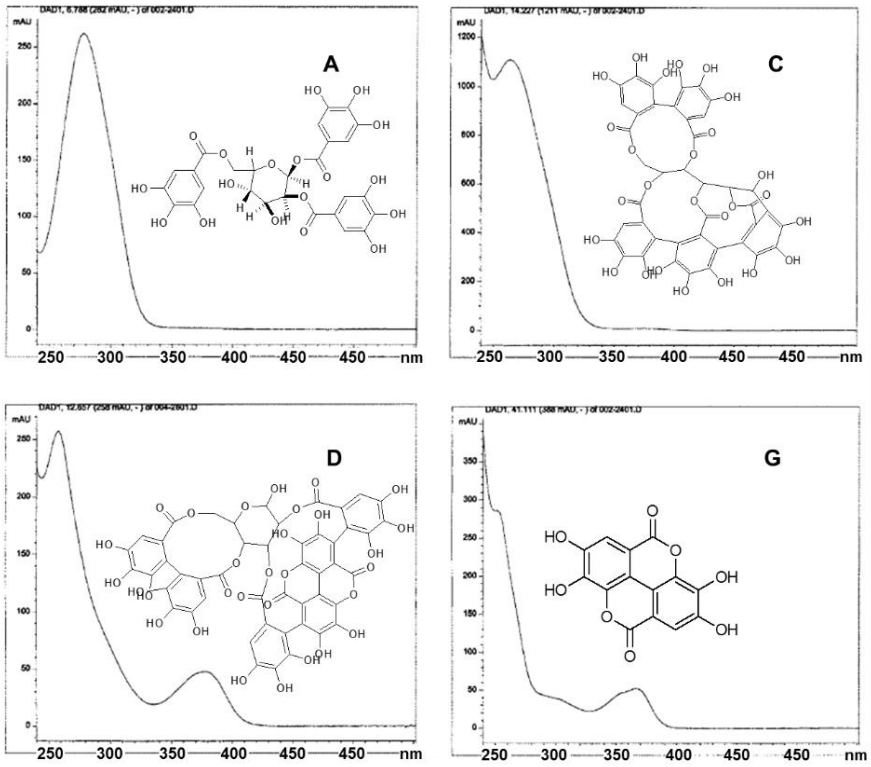


Figure 2.2: Representative UV spectra of different classes of hydrolysable tannins: **A**, gallotannin (trigalloyl glucose), **C** ellagitannin (vescalagin)**, D** ellagitannin (punicalagin) and **G** ellagic acid. Adapted from [60].

In addition to molecular weight information obtained from the molecular ion, fragmentation of the molecular ion can often provide useful information in the tentative identification of hydrolysable tannins [64]. For example, fragmentation of simple galloyl glucose species (gallotannins) typically involves neutral losses of a galloyl moieties (m/z 152). Furthermore, m/z 169 is a characteristic fragment ion identifying deprotonated gallic acid, which undergoes further decarboxylation (-CO_2, m/z 44) to produce the m/z 125 fragment ion [72]. On the other hand, fragmentation of

ellagitannins containing a HHDP unit produces the diagnostic fragment ions m/z 301 [ellagic acid-H]$^{1-}$, m/z 275 [ellagic acid-H-CO]$^{1-}$ and m/z 249 [ellagic acid-H-2CO]$^{1-}$ [64].

In recent years, considerable effort has been invested into the identification of the pomegranate phenolics. Pomegranates are a rich source of hydrolysable tannins which have been shown to exhibit *in vitro* antioxidant activity [60, 73–80]. For compound identification, these reports employed RP-LC on C_{18} columns using acetonitrile or methanol and acidified water (formic/acetic acid) as mobile phases and negative electrospray ionisation MS detection. In general, this methodology is employed for the separation and tentative identification of hydrolysable tannins in various foodstuffs [44].

An interesting analytical method was developed by Engström and co-workers [81], who reported a fast and selective method employing ultra-high pressure liquid chromatography (UHPLC) - triple quadrupole mass spectrometric (MS/MS) detection for the rapid fingerprinting of polyphenol subclasses in plants, including hydrolysable tannins [82]. This method utilises a high cone voltage followed by multiple reaction monitoring (MRM) data acquisition using common precursor and fragment ions to provide pseudo-MS^3 data for different compound classes, thereby increasing the selectivity of detection for fingerprinting purposes.

The question of the bioavailability of hydrolysable tannins has been the focus of numerous scientific papers, where biological samples such as urine, plasma and ileal fluids are typically analysed using RP-LC coupled to both UV and mass spectrometric detection [38–40, 83–85]. A bioavailability study conducted by Barrio *et al*. [40] demonstrated that negative atmospheric pressure chemical ionisation (APCI) provided lower limits of detection compared to electrospray ionisation (ESI) for urolithin metabolites in urine.

Matrix assisted laser desorption/ionisation (MALDI) has been widely employed in the characterisation of hydrolysable [13, 59, 68, 86–89] and condensed tannins [67, 90, 91]. Sáyago-Ayerdi *et al*. [86] investigated two different cationic agents, Na^+ and Cs^+, for the ionisation of hydrolysable tannins in mango peels using MALDI-TOF-MS. This allowed the detection of gallotannins from penta- (963 m/z) to trideca-O-galloylglucoses (m/z 2289), as confirmed by MALDI-TOF/TOF-MS fragmentation data.

MALDI-TOF was also used to probe the *in situ* macromolecular structure of hydrolysable tannins in chestnut wood. Pizzi and Pasch proposed a random tridimensional macromolecular network, formed by penta-O-galloylglucose monomers linked through a galloyl ether bond [87]. It was later proposed that the glucose units could be covalently bonded [68, 92].

In summary, for unambiguous structural elucidation of hydrolysable tannins, NMR is indispensable. However, for non-targeted analysis of complex mixtures of hydrolysable tannins, sample complexity and the low levels of occurrence of some constituents present obvious challenges, and LC-UV-MS provides a more realistic workflow for such applications. This is clearly evident from the number of research papers using primarily RP-LC-ESI-MS(/MS) for the analysis of poly(phenolics) in natural samples. In the following sections, two relatively recent analytical techniques which may be used to improve the performance of LC-MS workflows will be briefly described, as these will be explored for hydrolysable tannin analysis in this thesis.

2.2.1 Multidimensional chromatographic analysis

The performance of chromatographic separation methods may be compared using the concept of peak capacity (n_c), which is defined as the maximum number of compounds that can theoretically be resolved in a given separation space or time [93]. The peak capacity of a separation can be estimated using the following equation:

$$n_C = 1 + \frac{t_g}{\frac{1}{n}\sum_1^n w_b} \quad (1)$$

where t_g is the gradient time and w_b the average peak width at baseline, calculated for n number of peaks. Since peak capacity is a theoretical concept, no real-life separation will provide separation of the number of compounds equal to the peak capacity of the system, although the concept remains very useful in the comparison of ultimate separation performance of chromatographic methods. Indeed, based on statistical theory of component overlap, the peak capacity should exceed the number of compounds in the sample by a factor 100 to resolve 98% of randomly distributed peaks [94]. Owing to the structural diversity of polyphenolic compounds in plant material [95], the peak capacities obtainable using 1D LC (typically << 1000 n_c) clearly fall short of

being able to resolve these complex mixtures. One approach to overcome this limitation is to couple two (or more) independent separation modes. If this is done in a comprehensive manner (see below), the overall peak capacity of such a system ($n_{c,2D}$, Eq 2) would be the product of peak capacities in each separation dimension [96].

$$n_{c,2D} = {}^{1}n_c \times {}^{2}n_c \quad (2)$$

where ${}^{1}n_c$ and ${}^{2}n_c$ are the first- and second-dimension peak capacities, respectively. Multidimensional separations can be performed in either heart cutting or comprehensive modes [97]. The following discussion will focus on two-dimensional (2D) separations, as these are the most common and practical of multidimensional systems. In heart-cutting 2D chromatography, only specific fractions of the sample in the first dimension (^{1}D) are transferred to the second dimension (^{2}D) for further analysis. In contrast, in comprehensive 2D chromatography, the entire sample is analysed in both dimensions. For a 2D separation to be classified as comprehensive, it has to meet the following criteria (i) employ a different retention mechanism in each dimension (the requirement of orthogonality, see below), (ii) the separation achieved in the first dimension should essentially be preserved (which has implications for the sampling rate, see below), and (iii) equal percentages of all the components in the sample should be separated in both dimensions and detected [98].

Equation 2 represents a theoretical estimate of the maximum theoretical peak capacity of a comprehensive 2D system. This performance is rarely achieved in practice for two reasons: the finite orthogonality of the two separation steps, and under-sampling of the first-dimension separation. To evaluate the practical peak capacity ($n'_{c,2D}$) of a comprehensive 2D system, each of these aspects must be accounted for.

Orthogonality refers to the complementary information provided by each of the separation dimensions; for correlated separation systems, little additional separation is provided by their comprehensive hyphenation, and the resultant 2-dimensional separation space is not effectively utilised. The orthogonality of a 2D system may, therefore, be estimated by the degree of surface coverage by component peaks of the two-dimensional separation plane [99]. The fractional surface coverage (f_c) may be used to quantify the effective usage of separation space, and can be calculated using methods such as the convex hull approach [100]. This method uses an algorithm to

calculate the area of the smallest convex polygon that encloses all the peaks in the 2D separation space. The fractional surface coverage is then calculated by dividing the area of the convex polygon or 'hull' by the total area of the 2D separation.

Furthermore, to avoid the recombination of peaks that were partially separated in the first dimension, the sampling rate (i.e., the modulation period) should be fast enough to prevent what is referred to as "under-sampling" [101, 102]. The consensus is that each peak should be sampled at least 2-3 times to avoid a loss in resolution [102]. The extend of ^{1}D under-sampling can also be accounted for in the estimation of the practical peak capacity using an under-sampling correction factor, β [103], according to equation 3:

$$n^{*}_{c,2D} = \frac{{}^{1}n_c \times {}^{2}n_c \times f_c}{\beta} \quad (3)$$

where f_c is the fractional surface coverage and β is the under-sampling correction factor. Comprehensive two-dimensional chromatography can be performed using one of three different configurations, namely on-line, stop-flow and off-line [104]. The simplest of these is the off-line approach, where fractions of the ^{1}D effluent are collected and subsequently re-injected on the ^{2}D column. This implies that the two chromatographic separations are performed independently, and the ^{2}D analysis time is not subjected to restrictions based on the fraction collection time [104]. Although the off-line approach provides the highest peak capacities for this reason, the downside is the excessively long analysis times. In an on-line setup, eluent from the first dimension is continuously sampled using a modulator (switching valve) equipped with two sampling loops; a ^{1}D fraction is collected in one loop, while the content of the other is analysed in the ^{2}D. Fractions are collected for a period of time, the sampling time, that is equal to the total ^{2}D analysis or cycle time. This implies that the ^{2}D cycle time (the time required to complete the gradient and re-equilibrate the column) should be short enough to minimise ^{1}D under-sampling. Furthermore, during an on-line 2D separation, the mobile phase eluting from the ^{1}D becomes the injection solvent of the ^{2}D. In the case of two separation modes with incompatible mobile phases, such as is the case for hydrophilic interaction chromatography (HILIC) and RP-LC, modulation can lead to severe injection band broadening in the ^{2}D. To prevent this from happening, different approaches have been developed, which include (i) introducing a make-up flow which

dilutes the ^{1}D effluent with weak ^{2}D solvent [105], (ii) reducing the volume injected onto the second dimension column by using a flow splitter [104] and (iii) replacing the loops with trap columns [106]. Stop-flow two-dimensional chromatography is performed using an experimental setup which is similar to that used in on-line 2D chromatography, with the important difference that the ^{1}D flow is stopped following the transfer of a fraction for the duration of the ^{2}D separation of this fraction. This results in more time being available for the ^{2}D separation, and, therefore, higher peak capacities, but also analysis times similar to those of off-line 2D chromatographic methods [104].

While comprehensive two-dimensional LC (LC×LC) has found extensive application for the analysis of phenolics in general [107–109], only one paper utilising multidimensional LC for the analysis of hydrolysable tannins has been reported to date [110]. This study reported an on-line two-dimensional method for the separation of phenolics, including ellagitannins, in red raspberries using two RP-LC separations performed on different stationary phases coupled to UV and MS detection. Since very long sampling times (8 min) were used, this method doesn't meet the criteria to be considered a comprehensive 2D separation, but can rather be considered a multiple heart-cutting 2D LC separation [110].

2.2.2 Ion mobility spectrometry

Ion mobility (IM) spectrometry has in recent years experienced significant growth as an additional tool in combination with MS and LC, for the analysis of complex samples [111]. The separation mechanism in ion mobility spectrometry is based on the collision of gas phase ions with a neutral buffer gas while under the influence of an electric field. Under these conditions, larger or less compact ions will experience more collisions with the buffer gas, delaying their arrival times (drift times) compared to smaller or more compact ions with the same ionic charge [112]. The technique is, therefore, ideally compatible with MS detection, and provides a means of pre-filtering ions according to differences in their arrival times.

Different types of ion mobility spectrometers are available, including drift time ion mobility spectrometry (DTIMS) [113], travelling wave ion mobility spectrometry

(TWIMS) [114], field asymmetric waveform ion mobility spectrometry (FAIMS) [115] and trapped ion mobility spectrometry (TIMS) [116] instruments as well as a differential mobility analyser (DMA) [116]. The basic principles of each of these IM systems are graphically illustrated in **Figure 2.3** [111]**.** DTIMS and TWIMS are time dispersive instruments [117], where the arrival or drift time of an ion can be related to its rotationally averaged collisional cross section (CCS), thereby providing structural information about the ion. Time dispersive ion mobility spectrometers contain a drift tube comprised of many ring electrodes which serve as ion guides. DTIMS [113] employs a static uniform electric field to propel ions across the drift tube. In TWIMS, [114] a travelling voltage wave or 'travelling wave' is applied along the length of the RF ion guide which acts as drift tube for ion propulsion. Separation in TWIMS is caused by the ion either travelling ('surfing') along with this wave pulse, or 'rolling over' when the wave pulse overtakes the ion. Smaller ions will likely surf the wave with less roll over events, while larger ions will experience more roll-overs, thereby exiting the drift cell later. DTIMS systems operate under low field conditions, which allows for the direct determination of an ion's CCS value from the measured drift time. In contrast, determination of CCS values from measured ion arrival times in TWIMS requires calibration using a calibrant with known CCS values [118].

FAIMS and DMA instruments act as ion filters, preventing unwanted ions from entering the mass spectrometer, similar to the multiple reaction monitoring (MRM) acquisition mode used in triple quadrupole mass spectrometers. Thus, on these systems, not all ions can be observed under the same conditions, as is the case in DTIMS and TWIMS, and drift times are not recorded. Separation in FAIMS occurs by applying a suitable compensation voltage between two electrodes to allow only ions moving in the same direction as the carrier gas to exit the cell [115]. On the other hand, in DMA, ions move from one electrode to the other while an orthogonal flow of carrier gas only allows ions with specific mobilities to exit the cell [111]. Both FAIMS and DMS operate at atmospheric pressure.

TIMS is a recently developed technique which uses a high gas flow to push the ions forward with a force that is proportional to their area (CCS value), while at the same time an electric field is used to hold the ions in a stationary position. Ions are released

(separated) from the drift cell by decreasing the electric field, with the large ions exiting the drift cell first, followed by smaller ions [116].

Incorporating ion mobility spectrometry into LC-MS workflows has revealed several benefits of the technique for the analysis of complex samples. These include an increased duty cycle for low abundant compounds [119, 120], the option of filtering mass spectral data according to arrival time to simplify compound identification [121, 122], deconvolution of mass spectra comprising multiply charged species [121], separation of chromatographically co-eluting isomers and/or compounds [123], and a range of different fragmentation options [124].

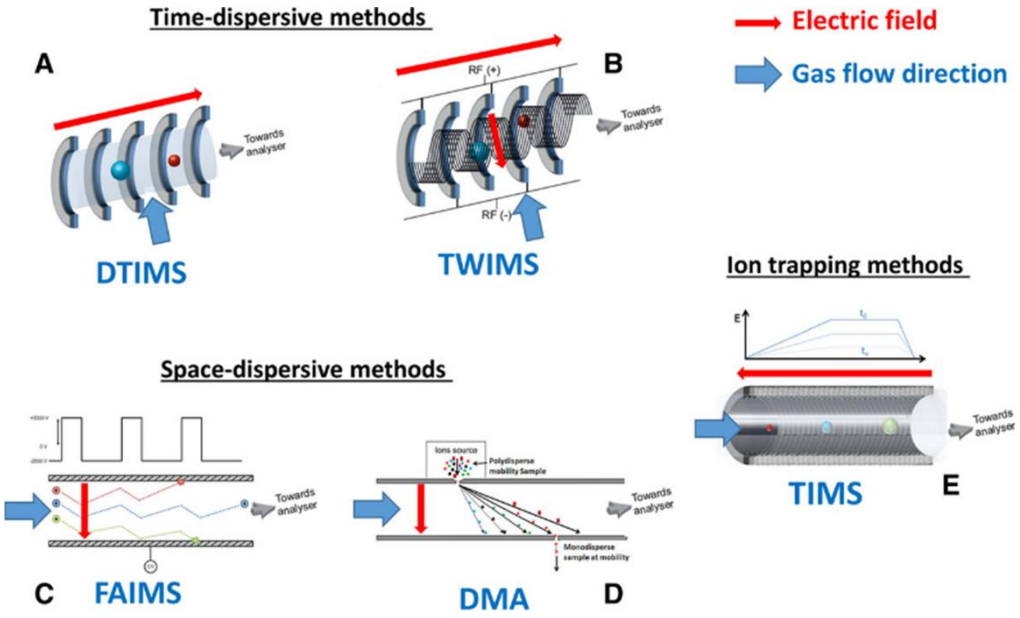


Figure 2.3: Graphical illustration of the different types of ion mobility spectrometers and their principles of operation. The instruments can be grouped into time-dispersive (DTIMS (**A**), and TWIMS (**B**)), space-dispersive methods (FAIMS (**C**), and DMA (**D**)) and ion trapping methods (TIMS (**E**)). Reprinted from [111].

The coupling of high performance liquid chromatographic (HPLC) separation to travelling wave ion mobility spectrometry-MS has in recent years revealed new information regarding the isomeric composition of polyphenols in a range of samples. Several papers on LC-IM-MS analysis of phenolics have reported the occurrence of

multiple IM peaks for a single chromatographic peak [125, 126]. This was observed for theasinensin C (a flavonoid dimer) [125] and eriodictyol glucopyranoside isomers (**Figure 2.4**) [126]. Due to the unique fragmentation spectra of catechin and epicatechin [127] in negative ionisation mode, Yassin *et al*. [125] postulated that the multiple drift times observed for theasinensin C could be due to *cis*-*trans*-isomerism at the C-2 position of the flavanol skeleton. Chalet *et al*. [128] showed that LC-TWIMS-MS provides valuable information regarding phase II metabolites of flavonols and flavones produced by HT29 (cancer) cells. This study revealed that ion mobility can differentiate between flavonoids glucuronidated at different positions, with 3′-O substituted compounds showing the smallest CCS values, followed by 3-O-glucuronides and 4′-O-glucuronides.

Similar multiple drift times observed for 5-O-caffeoyl quinic acid using TWIMS were initially hypothesised by Xie *et al*. [129] to be due to *cis* and *trans* isomers. However, Kuhnert *et al*. [130] later showed that the *cis/trans* 5-O-caffeoyl quinic acid geometric isomers could easily be separated on a reversed phase column, but that these nevertheless still produced multiple ion mobility peaks for each chromatographic peak. These authors [130], therefore, proposed that this phenomenon could be explained by prototropic isomerism, where the ionised caffeoyl quinic acid can exist as either a carboxylate or a phenolate anion. This hypothesis was supported by the presence of a single ion mobility peak for 5-O-caffeoyl quinic acid methylated at either the phenolic oxygen or at the carboxylate oxygen. Based on this work, the multiple arrival time distributions observed for the glycosylated eriodictyol species in rooibos tea were hypothesised to correspond to prototropic isomers [126] (**Figure 2.4**). In a recent study, the combination of ion mobility spectrometry with infrared spectroscopy provided direct evidence of multiple protonation sites on benzocaine, which were responsible for different drift times for each protomer [131].

Recently, several papers described the use of ESI-IM-MS to investigate the non-covalent interaction between polyphenols and proteins in the prevention of aggregation of specific proteins which are responsible for neurological disorders and type-2 diabetes mellitus [132–136]. The two polyphenols that have been evaluated for this purpose are the hydrolysable tannin pentagalloyl glucose and the procyanidin, epigallocatechin gallate. For example, a collision-induced unfolding (CIU)

methodology using a stepwise increase in the cone voltage was used to assess the strength of hydrogen bonds between superoxide dismutase (SOD1) and several catechins, including epigallocatechin gallate, epicatechin gallate and epigallocatechin [132]. Results showed that epigallocatechin gallate can maintain the compact state of COD1 at increased cone voltages and, therefore, shows promise as a potential candidate for further investigation.

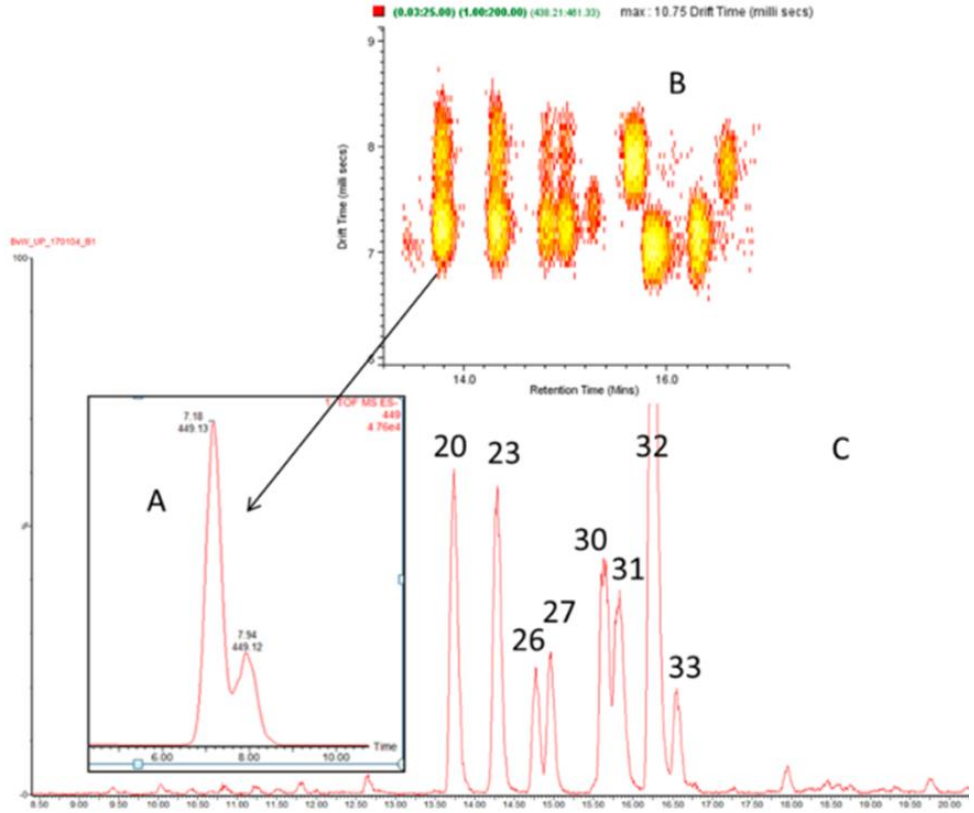


Figure 2.4: An extracted ion RP-LC×ion mobility contour plot obtained in the negative ion mode for m/z 449, 451 and 447 in rooibos tea. (**B**), illustrating multiple arrival times (**A**) for (S)-eriodictyol-6-*b*-D-glucopyranoside (**20**), (R)-eriodictyol-6-C-β-D-glucopyranoside (**23**), (S)-eriodictyol-8-C-β-D-glucopyranoside (**26**) and (R)-eriodictyol-8-C-β-D-glucopyranoside (**27**). Reprinted from [126].

Francheschi *et al*. [137] used TWIMS-MS to demonstrate that strong hydrogen bond interaction between ellagitannin oligomers is responsible for the formation of multiply charged clusters in ESI-MS, which were characterised by different arrival times.

Ion mobility as a stand-alone analytical technique in combination with MS is a rapidly developing field. Several papers have reported sub-1 minute analysis times using this approach to replace conventional LC-MS analyses. These include the use of DTIMS-MS for the rapid separation of selected isomeric flavonoids [138], food safety applications [139, 140], potency testing of cannabis [141] as well as monitoring of propofol levels in plasma to maintain anesthesia during surgical procedures [142]. Furthermore, stand-alone FAIMS has been employed for the rapid analysis (<1 min) of chlorogenic acids in orange juice [143], while TWIMS has been successfully used to separate (+)-catechin and (-)-epicatechin epimers using a suitable chiral ligand to enhance differences in their CCS values [144].

2.3 Summary

The comprehensive analysis of hydrolysable tannins is an important step in studying the biological roles that these compounds play in their native plants and in other living organisms. Furthermore, analysis of commercially available tannin extracts can provide better insight into their possible industrial applications. For the structural elucidation of hydrolysable tannins, NMR spectroscopy is essential. However, this requires milligram quantities of the pure compound, which requires elaborate extraction procedures and is often not feasible for trace-level constituents. Over the past few decades, the method of choice for the non-targeted analysis of hydrolysable tannins in plant material is RP-LC separation hyphenated to MS detection. Due to the molecular complexity of hydrolysable tannins, however, conventional one-dimensional LC methods often suffer from inadequate separation. In this context, two-dimensional liquid chromatography offers a means of improving the chromatographic separation, as has been demonstrated for various poly(phenolics), but not yet for hydrolysable tannins. Furthermore, the incorporation of ion mobility spectrometry into the LC-MS workflow could potentially further improve data quality and the performance of these methods.

2.4 References

1. Quideau S, Deffieux D, Douat-Casassus C, Pouysegu L (2011) Natural Products Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. Angew Chem 50:586–621. doi: 10.1002/anie.201000044

2. Okuda T, Yoshida T, Hatano T (2000) Correlation of oxidative transformations of hydrolyzable tannins and plant evolution. Phytochemistry 55:513–529. doi: 10.1016/S0031-9422(00)00232-6

3. Jurd L (1958) Plant Polyphenols. III. The Isolation of a New Ellagitannin from Pellicle of the Walnut. J Am Chem Soc 80:2249–2252

4. White T (1957) Tannins-Their occurrence and significance. J Sci Food Agric 8:377–385. doi: 10.1002/jsfa.2740080702

5. Tuominen A, Sundman T (2013) Stability and oxidation products of hydrolysable tannins in basic conditions detected by HPLC/DAD-ESI/QTOF/MS. Phytochem Anal 24:424–435. doi: 10.1002/pca.2456

6. Haslam E, Cai Y (1994) Plant polyphenols (vegetable tannins): Gallic acid metabolism. Nat Prod Rep 11:41–66. doi: 10.1039/np9941100041

7. Schmidt OJ, Mayer W (1956) Natürliche Gerbstoffe. Angew Chem 68:103–115

8. Herve du Penhoat CLM, Michon VMF, Ohassan A, Peng S, Scalbert A, Gage D (1991) Roburin A, A dimeric ellagitannin from heartwood of Quercus robur. Phytochemistry 30:329–332. doi: 10.1016/0031-9422(91)84148-L

9. Quideau S, Jourdes M, Lefeuvre D, Montaudon D, Saucier C, Glories Y, Pardon P, Pourquier P (2005) The chemistry of wine polyphenolic C-glycosidic ellagitannins targeting human topoisomerase II. Chem - A Eur J 11:6503–6513. doi: 10.1002/chem.200500428

10. Okuda T, Yoshida T, Hatano T (1990) Oligomeric hydrolysable tannins, a new class of plant polyphenols. Heterocycles 30:1195–1218

11. Spina S, Zhou X, Segovia C, Pizzi A, Romagnoli M, Giovando S, Pasch H, Rode K, Delmotte L (2013) Phenolic resin adhesives based on chestnut (Castanea

sativa) hydrolysable tannins. J Adhes Sci Technol 27:2103–2111. doi: 10.1080/01694243.2012.697673

12. Aouf C, Benyahya S, Esnouf A, Caillol S, Boutevin B, Fulcrand H (2014) Tara tannins as phenolic precursors of thermosetting epoxy resins. Eur Polym J 55:186–198

13. Abdalla S, Pizzi A, Bahabri F, Ganash A (2015) Analysis of Valonia Oak (Quercus aegylops) Acorn Tannin and Wood Adhesives Application. Bioresources 10:7165–7177

14. Ghahri S, Pizzi A, Mohebby B, Mirshokraie A, Mansouri HR (2018) Soy-based, tannin-modified plywood adhesives. J Adhes 94:218–237. doi: 10.1080/00218464.2016.1258310

15. Lagel MC, Pizzi A, Giovando S (2014) Matrix-Assisted Laser Desorption-Ionization Time of Flight (MALDI-TOF) Mass Spectrometry of Phenol-Formaldehyde-Chestnut Tannin Resins. J Renew Mater 2:207–219

16. Sanz M, Cadahia E, Esteruelas E, Munoz AM, Fernandez De Simon B, Hernandez T, Estrella I, Cadahia E, Esteruelas E, Munoz AM, Fernandez De Simon B, Hernandez T, Estrella I (2010) Phenolic compounds in chestnut (Castanea sativa Mill.) heartwood. Effect of toasting at cooperage. J Agric Food Chem 58:9631–9640. doi: 10.1021/jf102718t

17. Merino SF, Caprari JJ, Torres LV, Ramos LF, Girola AH (2017) Inhibitive action of tara tannin in rust converter formultaion. Anti-Corrosion Methods Mater 64:136–147

18. Appel H (1993) Phenolics in ecological interactions: The importance of oxidation. J Chem Ecol 19:1521–1552

19. Feeny PP (1969) Inhibitory effect of oak tannins on the hydrolysis of proteins by trypsin. Phytochemistry 8:2119–2126

20. Feeny PP (1968) Effect of oak leaf tannins on larval growth of tghe winter moth Operophtera Brumata. J Insect Physiol 14:805–817

21. Moilanen J, Karonen M, Tähtinen P, Jacquet R, Quideau S, Salminen J (2016) Phytochemistry Biological activity of ellagitannins: Effects as anti-oxidants, pro-oxidants and metal chelators. Phytochemistry 125:65–72. doi: 10.1016/j.phytochem.2016.02.008

22. Gyamfi MA, Aniya Y (2002) Antioxidant properties of Thonningianin A, isolated from the African medicinal herb, Thonningia sanguinea. Biochem Pharmacol 63:1725–1737. doi: 10.1016/S0006-2952(02)00915-2

23. Heilman J, Andreux P, Tran N, Rinsch C, Blanco-Bose W (2017) Safety assessment of Urolithin A, a metabolite produced by the human gut microbiota upon dietary intake of plant derived ellagitannins and ellagic acid. Food Chem Toxicol 108:289–297. doi: 10.1016/j.fct.2017.07.050

24. Cerdá B, Espín JC, Parra S, Martínez P, Tomás-Barberán FA (2004) The potent in vitro antioxidant ellagitannins from pomegranate juice are metabolised into bioavailable but poor antioxidant hydroxy-6H-dibenzopyran-6-one derivatives by the colonic microflora of healthy humans. Eur J Nutr 43:205–220. doi: 10.1007/s00394-004-0461-7

25. Larrosa M, García-Conesa MT, Espín JC, Tomás-barberán FA (2010) Molecular Aspects of Medicine Ellagitannins, ellagic acid and vascular health. Mol Aspects Med 31:513–539. doi: 10.1016/j.mam.2010.09.005

26. Päivärinta E, Pajari A, Törrönen R, Mutanen M (2009) Ellagic Acid and Natural Sources of Ellagitannins as Possible Chemopreventive Agents Against Intestinal Tumorigenesis in the Min Mouse. Nutr Cancer 54:79–83. doi: 10.1207/s15327914nc5401

27. Larrosa M, Tomas-Barberan FA, Espin JC (2006) The dietary hydrolysable tannin punicalagin releases ellagic acid that induces apoptosis in human colon adenocarcinoma Caco-2 cells by using the mitochondrial pathway. J Nutr Biochem 17:611–625. doi: 10.1016/j.jnutbio.2005.09.004

28. Xu J, Yuan C, Wang G, Luo J, Xu L, Mu Y, Li Y, Seeram NP, Huang X, Li L (2018) Urolithins Attenuate LPS-Induced Neuroinflammation in BV2Microglia via MAPK, Akt, and NF-κB Signaling Pathways. J Agric Food Chem 66:571–580.

doi: 10.1021/acs.jafc.7b03285

29. Tomas-Barberan FA, Gonzalez-Sarrıas A, Garcıa-Villalba R, Nunez-Sanchez MA, Selma MV, Garcıa-Conesa MT, Espın JC (2017) Urolithins, the rescue of “old” metabolites to understand a “new” concept: Metabotypes as a nexus among phenolic metabolism, microbiota dysbiosis, and host. Mol Nutr Food Res 61:1–35. doi: 10.1002/mnfr.201500901

30. Auzanneau C, Montaudon D, Jacquet R, Puyo S, Pouységu L, Deffieux D, Elkaoukabi-chaibi A, De Giorgi F, Ichas F, Quideau S, Pourquier P (2012) The Polyphenolic Ellagitannin Vescalagin Acts As a Preferential Catalytic Inhibitor of the α Isoform of Human DNA Topoisomerase II. Mol Pharmacol 82:134–141

31. Hooper L, Kroon PA, Rimm EB, Cohn JS, Harvey I, Le Cornu KA, Ryder JJ, Hall WL, Cassidy A (2008) Flavonoids, flavonoid-rich foods, and cardiovascular risk: a meta-analysis of randomized controlled trials. Am J Clin Nutr 88:38–50. doi: 10.1093/ajcn/88.1.38

32. Sylla T, Pouységu L, Da Costa G, Deffieux D, Monti J, Quideau S (2015) Gallotannins and Tannic Acid: First Chemical Syntheses and In Vitro Inhibitory Activity on Alzheimer’s Amyloid β -Peptide Aggregation. Angew Chem 54:8217–8221. doi: 10.1002/anie.201411606

33. Sosic A, Cappellini M, Sinigaglia L, Jacquet R, Deffieux D, Fabris D, Quideau S, Gatto B, Fabris D (2015) Polyphenolic C-glucosidic ellagitannins present in oak-aged wine inhibit HIV-1 nucleocapsid protein. Tetrahedron 71:3020–3026. doi: 10.1016/j.tet.2015.01.035

34. Vilhelmova-ilieva N, Jacquet R, Quideau S, Galabov AS (2014) Ellagitannins as synergists of ACV on the replication of ACV-resistant strains of HSV 1 and 2. Antiviral Res 110:104–114. doi: 10.1016/j.antiviral.2014.07.017

35. Bae S, Kim SY, Do MH, Lee CH, Song YJ (2017) 1,2,3,4,6-Penta-O-galloyl-ß-D-glucose, a bioactive compound in Elaeocarpus sylvestris extract, inhibits varicella-zoster virus replication. Antiviral Res 144:266–272. doi: 10.1016/j.antiviral.2017.06.018

36. Pizzi A (2008) Tannins: Major Sources, Properties and Applications, Chapter 8. In: Monomers, Polymers and Composites from Renewable Resources. pp 179–200

37. Trinh BTD, Staerk D, Jäger AK (2016) Screening for potential α-glucosidase and α-amylase inhibitory constituents from selected Vietnamese plants used to treat type 2 diabetes. J Ethnopharmacol 186:189–195. doi: 10.1016/j.jep.2016.03.060

38. Cerdá B, Llorach R, Cerón JJ, Espín JC, Tomás-Barberán FA (2003) Evaluation of the bioavailability and metabolism in the rat of punicalagin, an antioxidant polyphenol from pomegranate juice. Eur J Nutr 42:18–28. doi: 10.1007/s00394-003-0396-4

39. Mertens-Talcott SU, Jilma-Stohlawetz P, Rios J, Hingorani L, Derendorf H (2006) Absorption, Metabolism, and Antioxidant Effects of Pomegranate (Punica granatum L.) Polyphenols after Ingestion of a Standardized Extract in Healthy Human Volunteers. J Agric Food Chem 54:8956–8961. doi: 10.1021/jf061674h

40. Gonzales-Barrio R, Borges G, Mullen W, Grozier A (2010) Bioavailability of anthocyanins and ellagitannins following consumption of raspberries by healthy humans and subjects. J Agric Food Chem 58:3933–3939. doi: 10.1021/jf100315d

41. Selma VM, Beltran D, Garcıa-Villalba R, Espin JC, Tomás-barberán FA (2014) Description of urolithin production capacity from ellagic acid of two human intestinal Gordonibacter species. Food Funct 5:1779–1784. doi: 10.1039/c4fo00092g

42. Cerdá B, Tomás-Barberán FA, Espin JC (2005) Metabolism of Antioxidant and Chemopreventive Ellagitannins from Strawberries, Raspberries, Walnuts, and Oak-Aged Wine in Humans: Identification of Biomarkers and Individual Variability. J Agric Food Chem 53:227–235. doi: 10.1021/jf049144d

43. Cai Y, Zhang J, Chen NG, Shi Z, Qiu J, He C, Chen M (2017) Recent Advances in Anticancer Activities and Drug Delivery Systems of Tannins. Med Res Rev 37:665–701. doi: 10.1002/med

44. Arapitsas P (2012) Hydrolyzable tannin analysis in food. Food Chem 135:1708–1717. doi: 10.1016/j.foodchem.2012.05.096

45. Tian F, Li B, Ji B, Yang J, Zhang G, Chen Y, Luo Y (2009) Antioxidant and antimicrobial activities of consecutive extracts from Galla chinensis: The polarity affects the bioactivities. Food Chem 113:173–179. doi: 10.1016/j.foodchem.2008.07.062

46. Tian F, Li B, Ji B, Zhang G, Luo Y (2009) Identification and structure-activity relationship of gallotannins separated from Galla chinensis. LWT - Food Sci Technol 42:1289–1295. doi: 10.1016/j.lwt.2009.03.004

47. Singh RP, Chidambara Murthy KN, Jayaprakasha GK (2002) Studies on the antioxidant activity of pomegranate (Punica granatum) peel and seed extracts using in vitro models. J Agric Food Chem 50:81–86. doi: 10.1021/jf010865b

48. Çam M, Hişil Y (2010) Pressurised water extraction of polyphenols from pomegranate peels. Food Chem 123:878–885. doi: 10.1016/j.foodchem.2010.05.011

49. Okuda T, Mori K, Seno K, Hatano T (1979) Constituents of Geranium thunbergII Sieb. et Zucc. J Chromatogr 171:313–320. doi: 10.1016/S0021-9673(01)95310-5

50. Okuda T, Yoshida T, Hatano T, Yazaki K, Ashida M (1980) Ellagitannins of the casuarinaceae, stachyuraceae and myrtaceae. Phytochemistry 21:2871–2874. doi: 10.1016/0031-9422(80)85058-8

51. Okuda T, Yoshida T, Hatano T, Yazaki K, Kira R, Ikeda Y (1986) Chromatography of Tannins II. Preparative Fractionation of Hydrolyzable Tannins By Centrifugal Partition Chromatography. J Chromatogr chromatograhy 362:375–381

52. Hatano T, Yoshida T, Okuda T (1988) Multiple peaks in high performance liquid chromatography of some hydrolyzable tannins. J Chromatogr 435:285–295

53. Scalbert A, Duval L, Peng S, Monties B, Du Penhoat C (1990) Polyphenols of Quercus robur L. II[a]. Preparative isolation by low-pressure and high-pressure

liquid chromatography of heartwood ellagitannins. J Chromatogr 502:107–119

54. Navarro M, Kontoudakis N, Miquel J, García-Romero E, Gómez-Alonso S, Zamora F, Hermosín-Gutiérrez I (2017) Improved method for the extraction and chromatographic analysis on a fused-core column of ellagitannins found in oak-aged wine. Food Chem 226:23–31. doi: 10.1016/j.foodchem.2017.01.043

55. Glabasnia A, Hofmann T (2006) Sensory-Directed Identification of Taste-Active Ellagitannins in American (Quercus alba L.) and European Oak Wood (Quercus robur L.) and Quantitative Analysis in Bourbon Whiskey and Oak-Mutured Red Wines. J Agric Food Chem 54:3380–3390. doi: 10.1021/jf052617b

56. Suvanto J, Tähtinen P, Valkamaa S, Engström MT, Karonen M, Salminen JP (2018) Variability in Foliar Ellagitannins of Hippophaë rhamnoides L. and Identification of a New Ellagitannin, Hippophaenin C. J Agric Food Chem 66:613–620. doi: 10.1021/acs.jafc.7b04834

57. Piwowarski JP, Kiss AK (2013) C-glucosidic ellagitannins from lythri herba (European Pharmacopoeia): Chromatographic profile and structure determination. Phytochem Anal 24:336–348. doi: 10.1002/pca.2415

58. Engels C, Knödler M, Zhao YY, Carle R, Gänzle MG, Schieber A (2009) Antimicrobial activity of gallotannins isolated from mango (Mangifera indica L.) kernels. J Agric Food Chem 57:7712–7718. doi: 10.1021/jf901621m

59. Kool MM, Comeskey DJ, Cooney JM, McGhie TK (2010) Structural identification of the main ellagitannins of a boysenberry (Rubus loganbaccus × baileyanus Britt.) extract by LC-ESI-MS/MS, MALDI-TOF-MS and NMR spectroscopy. Food Chem 119:1535–1543. doi: 10.1016/j.foodchem.2009.09.039

60. Gil MI, Tomas-Barberan FA, Hess-Pierce B, Holcroft DM, Kader AA (2000) Antioxidant Activity of Pomegranate Juice and Its Relationship with Phenolic Composition and Processing. J Agric Food Chem 48:4581–4589. doi: 10.1021/jf000404a

61. Moilanen J, Sinkkonen J, Salminen JP (2013) Characterization of bioactive plant ellagitannins by chromatographic, spectroscopic and mass spectrometric

methods. Chemoecology 23:165–179. doi: 10.1007/s00049-013-0132-3

62. Salminen JP, Ossipov V, Loponen J, Haukioja E, Pihlaja K (1999) Characterisation of hydrolysable tannins from leaves of Betula pubescens by high-performance liquid chromatography-mass spectrometry. J Chromatogr A 864:283–291. doi: 10.1016/S0021-9673(99)01036-5

63. Jurd L (1956) Plant Polyphenols. I. The Polyphenolic Constituents of the Pellicle of the Walnut. J Am Chem Soc 78:3345–3448. doi: 10.1021/ja01595a050

64. Arapitsas P, Menichetti S, Vincieri FF, Romani A (2007) Hydrolyzable tannins with the hexahydroxydiphenoyl unit and the m-depsidic link: HPLC-DAD-MS identification and model synthesis. J Agric Food Chem 55:48–55

65. Regueiro J, Sánchez-González C, Vallverdú-Queralt A, Simal-Gándara J, Lamuela-Raventós R, Izquierdo-Pulido M (2014) Comprehensive identification of walnut polyphenols by liquid chromatography coupled to linear ion trap – Orbitrap mass spectrometry. Food Chem 152:340–348. doi: 10.1016/j.foodchem.2013.11.158

66. Afaq F, Saleem M, Krueger CG, Reed JD, Mukhtar H (2005) Anthocyanin- and hydrolyzable tannin-rich pomegranate fruit extract modulates MAPK and NF-κB pathways and inhibits skin tumorigenesis in CD-1 mice. Int J Cancer 113:423–433. doi: 10.1002/ijc.20587

67. Saad H, Bouhtoury FC, Pizzi A, Rode K, Charrier B, Ayed N (2012) Characterization of pomegranate peels tannin extractives. Ind Crop Prod 40:239–246. doi: 10.1016/j.indcrop.2012.02.038

68. Pizzi, A. Pasch, H. Rode, K. Giovando S (2009) Polymer Structure of Commercial Hydrolyzable Tanins by Matrix-Assisted Laser Desorption/Ionization-Time-of-Flight Mass Spectriometry. J Appl Polym Sci 113:3847–3859

69. Xiang P, Lin Y, Lin P, Xiang C, Yang Z, Lu Z (2007) Effect of cationization reagents on the matrix-assisted laser desorption/ionization time-of-flight mass spectrum of chinese gallotannins. J Appl Polym Sci 105:859–864. doi:

10.1002/app.26373

70. Rauha JP, Vuorela H, Kostiainen R (2001) Effect of eluent on the ionization efficiency of flavonoids by ion spray, atmospheric pressure chemical ionization, and atmospheric pressure photoionization mass spectrometry. J Mass Spectrom 36:1269–1280. doi: 10.1002/jms.231

71. Domon B, Hostettmann K (1985) Mass spectrometric studies of underivatized polyphenolic glycosides. Phytochemistry 24:575–580. doi: 10.1016/S0031-9422(00)80769-4

72. Meyers KJ, Swiecki TJ, Mitchell AE (2006) Understanding the Native Californian diet: Identification of condensed and hydrolyzable tannins in tanoak acorns (Lithocarpus densiflorus). J Agric Food Chem 54:7686–7691. doi: 10.1021/jf061264t

73. Mena P, Calani L, Dall'Asta C, Galaverna G, García-Viguera C, Bruni R, Crozier A, Del Rio D (2012) Rapid and comprehensive evaluation of (Poly)phenolic compounds in pomegranate (Punica granatum L.) Juice by UHPLC-MSn. Molecules 17:14821–14840. doi: 10.3390/molecules171214821

74. Legua P, Forner-Giner MA, Nuncio-Jáuregui N, Hernández F (2016) Polyphenolic compounds, anthocyanins and antioxidant activity of nineteen pomegranate fruits: A rich source of bioactive compounds. J Funct Foods 23:628–636. doi: 10.1016/j.jff.2016.01.043

75. Ambigaipalan P, de Camargo AC, Shahidi F (2016) Phenolic Compounds of Pomegranate Byproducts (Outer Skin, Mesocarp, Divider Membrane) and Their Antioxidant Activities. J Agric Food Chem 64:6584–6604. doi: 10.1021/acs.jafc.6b02950

76. Cano-Lamadrid M, Marhuenda-Egea FC, Hernández F, Rosas-Burgos EC, Burgos-Hernández A, Carbonell-Barrachina AA (2016) Biological Activity of Conventional and Organic Pomegranate Juices: Antioxidant and Antimutagenic Potential. Plant Foods Hum Nutr 71:375–380. doi: 10.1007/s11130-016-0569-y

77. Ambigaipalan P, de Camargo AC, Shahidi F (2017) Identification of phenolic

antioxidants and bioactives of pomegranate seeds following juice extraction using HPLC-DAD-ESI-MSn. Food Chem 221:1883–1894. doi: 10.1016/j.foodchem.2016.10.058

78. Abdulla R, Mansur S, Lai H, Ubul A, Sun G, Huang G, Aisa HA (2017) Qualitative Analysis of Polyphenols in Macroporous Resin Pretreated Pomegranate Husk Extract by HPLC-QTOF-MS. Phytochem Anal 28:465–473. doi: 10.1002/pca.2695

79. Aguilar-Zárate P, Wong-Paz JE, Michel M, Buenrostro-Figueroa J, Díaz HR, Ascacio JA, Contreras-Esquivel JC, Gutiérrez-Sánchez G, Aguilar CN (2017) Characterisation of Pomegranate-Husk Polyphenols and Semi-Preparative Fractionation of Punicalagin. Phytochem Anal 28:433–438. doi: 10.1002/pca.2691

80. Russo M, Fanali C, Tripodo G, Dugo P, Muleo R, Dugo L, De Gara L, Mondello L (2018) Analysis of phenolic compounds in different parts of pomegranate (Punica granatum) fruit by HPLC-PDA-ESI/MS and evaluation of their antioxidant activity: application to different Italian varieties. Anal Bioanal Chem 410:3507–3520. doi: 10.1007/s00216-018-0854-8

81. Engström MT, Pälijärvi M, Salminen JP (2015) Rapid fingerprint analysis of plant extracts for ellagitannins, gallic acid, and quinic acid derivatives and quercetin-, kaempferol- and myricetin-based flavonol glycosides by UPLC-QqQ-MS/MS. J Agric Food Chem 63:4068–4079. doi: 10.1021/acs.jafc.5b00595

82. Salminen J (2018) Two-Dimensional Tannin Fingerprints by Liquid Chromatography Tandem Mass Spectrometry Offer a New Dimension to Plant Tannin Analyses and Help To Visualize the Tannin Diversity in Plants. J Agric Food Chem 66:9162–9171. doi: 10.1021/acs.jafc.8b02115

83. Seeram NP, Lee R, Heber D (2004) Bioavailability of ellagic acid in human plasma after consumption of ellagitannins from pomegranate (Punica granatum L.) juice. Clin Chim Acta 348:63–68. doi: 10.1016/j.cccn.2004.04.029

84. Seeram NP, Henning SM, Zhang Y, Suchard M, Li Z, Heber D (2006) Pomegranate juice ellagitannin metabolites are present in human plasma and

some persist in urine for up to 48 hours. J Nutr 136:2481–2485

85. García-Villalba R, Espín JC, Tomás-Barberán FA (2016) Chromatographic and spectroscopic characterization of urolithins for their determination in biological samples after the intake of foods containing ellagitannins and ellagic acid. J Chromatogr A 1428:162–175. doi: 10.1016/j.chroma.2015.08.044

86. Sáyago-Ayerdi S, Moreno-Hernández C, Montalvo-González E, García-Magaña M, Lluís J, Pérez-Jiménez J (2013) Mexican 'Ataulfo' mango (Mangifera indica L) as a source of hydrolyzable tannins. Analysis by MALDI-TOF/TOF MS. Food Res Int 51:188–194. doi: 10.1016/j.foodres.2012.11.034

87. Pasch H, Pizzi A (2002) Considerations on the macromolecular structure of chestnut ellagitannins by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. J Appl Polym Sci 85:429–437. doi: 10.1002/app.10618

88. Radebe N, Rode K, Pizzi A, Pasch H (2013) Microstructure elucidation of polyflavonoid tannins by MALDI-TOF-CID. J Appl Polym Sci 127:1937–1950. doi: 10.1002/app.37568

89. Reed JD, Krueger CG, Vestling MM (2013) MALDI-TOF mass spectrometry of oligomeric food polyphenols. Int Food Res J 66:2248–2263. doi: 10.1016/j.phytochem.2005.05.015

90. Mahmoud SB, Saad H, Charrier B, Pizzi A, Rode K, Ayed N, Bouhtoury FC (2015) Characterization of sumac (Rhus tripartitum) root barks tannin for a potential use in wood adhesives formulation. Wood Sci Technol 49:205–221. doi: 10.1007/s00226-014-0686-4

91. Pasch H, Pizzi A, Rode K (2001) MALDI-TOF mass spectrometry of polyflavonoid tannins. Polymer (Guildf) 42:7531–7539. doi: 10.1016/S0032-3861(01)00216-6

92. Radebe N, Rode K, Pizzi A, Giovando S, Pasch H (2013) MALDI-TOF-CID for the microstructure elucidation of polymeric hydrolysable tannins. J Appl Polym Sci 128:97–107. doi: 10.1002/app.38156

93. Grushka E (1970) Chromatographic Peak Capacity and the Factors Influencing

It. Anal Chem 42:1142–1147. doi: 10.1021/ac60293a001

94. Davis JM, Giddings JC (1983) Statistical-Theory of Component Overlap in Multicomponent Chromatograms. Anal Chem 55:418–424. doi: 10.1021/ac00254a003

95. Haslam E (2007) Vegetable tannins - Lessons of a phytochemical lifetime. Phytochemistry 68:2713–2721. doi: 10.1016/j.phytochem.2007.09.009

96. Giddings JC (1984) Two-dimensional separations: concept and promise. Anal Chem 56:1258A–1270A. doi: 10.1021/ac00276a003

97. Schure MR (1999) Limit of detection, dilution factors, and technique compatibility in multidimensional separations utilizing chromatography, capillary electrophoresis, and field-flow fractionation. Anal Chem 71:1645–1657. doi: 10.1021/ac981128q

98. Schoenmakers PJ, Marriot P, Beens J (2003) Nomenclature and conventions in comprehensive multidimensional chromatography. J Chromatogr A 16:335–339

99. Gilar M, Olivova P, Daly AE, Gebler C (2005) Orthogonality of Separation in Two-Dimensional Liquid Chromatography. Anal Chem 77:6426–6434. doi: 10.1021/ac050923i

100. Semard G, Peulon-Agasse V, Bruchet A, Bouillon JP, Cardinaël P (2010) Convex hull: A new method to determine the separation space used and to optimize operating conditions for comprehensive two-dimensional gas chromatography. J Chromatogr A 1217:5449–5454. doi: 10.1016/j.chroma.2010.06.048

101. Gu H, Huang Y, Carr PW (2011) Peak capacity optimization in comprehensive two dimensional liquid chromatography: A practical approach. J Chromatogr A 1218:64–73. doi: 10.1016/j.chroma.2010.10.096

102. Murphy RE, Schure MR, Foley JP (1998) Effect of Sampling Rate on Resolution in Comprehensive Two-Dimensional Liquid Chromatography. Anal Chem 70:1585–1594. doi: 10.1021/ac971184b

103. Li X, Stoll DR, Carr PW (2009) A Simple and Accurate Equation for Peak Capacity Estimation in Two Dimensional Liquid Chromatography. Anal Chem 81:845–850. doi: 10.1021/ac801772u

104. Kalili KM, De Villiers A (2013) Systematic optimisation and evaluation of on-line, off-line and stop-flow comprehensive hydrophilic interaction chromatography × reversed phase liquid chromatographic analysis of procyanidins. Part II: Application to cocoa procyanidins. J Chromatogr A 1289:69–79. doi: 10.1016/j.chroma.2013.03.009

105. Stoll DR, Talus ES, Harmes DC (2015) Evaluation of detection sensitivity in comprehensive two-dimensional liquid chromatography separations of an active pharmaceutical ingredient and its degradants. Anal Bioanal Chem 407:265–277. doi: 10.1007/s00216-014-8036-9

106. Li Q, Lynen F, Wang J, Li H, Xu G, Sandra P (2012) Comprehensive hydrophilic interaction and ion-pair reversed-phase liquid chromatography for analysis of di- to deca-oligonucleotides. J Chromatogr A 1255:237–243. doi: 10.1016/j.chroma.2011.11.062

107. Villiers A de, Venter P, Pasch H (2016) Recent advances and trends in the liquid-chromatography – mass spectrometry analysis of flavonoids. J Chromatogr A 1430:16–78. doi: 10.1016/j.chroma.2015.11.077

108. Willemse CM, Stander MA, Vestner J, Tredoux AGJ, de Villiers A (2015) Comprehensive Two-Dimensional Hydrophilic Interaction Chromatography (HILIC) × Reversed-Phase Liquid Chromatography Coupled to High-Resolution Mass Spectrometry (RP-LC-UV-MS) Analysis of Anthocyanins and Derived Pigments in Red Wine. Anal Chem 87:12006–12015. doi: 10.1021/acs.analchem.5b03615

109. Cacciola F, Farnetti S, Dugo P, Marriott PJ, Mondello L (2017) Comprehensive two-dimensional liquid chromatography for polyphenol analysis in foodstuffs. J Sep Sci 40:7–24. doi: 10.1002/jssc.201600704

110. Kula M, Głód D, Krauze-Baranowska M (2016) Two-dimensional liquid chromatography (LC) of phenolic compounds from the shoots of Rubus idaeus

"Glen Ample" cultivar variety. J Pharm Biomed Anal 121:99–106. doi: 10.1016/j.jpba.2015.12.047

111. D'Atri V, Causon T, Hernandez-Alba O, Mutabazi A, Veuthey JL, Cianferani S, Guillarme D (2018) Adding a new separation dimension to MS and LC–MS: What is the utility of ion mobility spectrometry? J Sep Sci 41:20–67. doi: 10.1002/jssc.201700919

112. Clemmer DE, Jarrold MF (1997) Ion mobility measurement and their applications to cluster and biomolecules. J Mass Spectrom 32:577–592

113. Wyttenbach T, Kemper PR, Bowers MT (2001) Design of a new electrospray ion mobility mass spectrometer. Int J Mass Spectrom 212:13–23

114. Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH (2004) Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. Rapid Commun Mass Spectrom 18:2401–2414. doi: 10.1002/rcm.1641

115. Guevremont R (2004) High-field asymmetric waveform ion mobility spectrometry: A new tool for mass spectrometry. J Chromatogr A 1058:3–19. doi: 10.1016/j.chroma.2004.08.119

116. Michelmann K, Silveira JA, Ridgeway ME, Park MA (2015) Fundamentals of trapped ion mobility spectrometry. J Am Soc Mass Spectrom 26:14–24. doi: 10.1007/s13361-014-0999-4

117. May JC, Mclean JA (2015) Ion Mobility-Mass Spectrometry: Time-Dispersive Instrumentation. Anal Chem. doi: 10.1021/ac504720m

118. Forsythe JG, Petrov AS, Walker CA, Allen SJ, Pellissier JS, Bush MF, Hud N V., Fernández FM (2015) Collision cross section calibrants for negative ion mode traveling wave ion mobility-mass spectrometry. Analyst 140:6853–6861

119. Zheng X, Wojcik R, Zhang X, Ibrahim YM, Burnum-Johnson KE, Orton DJ, Monroe ME, Moore RJ, Smith RD, Baker ES (2017) Coupling Front-End Separations, Ion Mobility Spectrometry, and Mass Spectrometry For Enhanced Multidimensional Biological and Environmental Analyses. Annu Rev Anal Chem

10:71–92. doi: 10.1146/annurev-anchem-061516-045212

120. Venne K, Bonneil E, Eng K, Thibault P (2005) Improvement in Peptide Detection for Proteomics Analyses Using NanoLC-MS and High-Field Asymmetry Waveform Ion Mobility Mass Spectrometry. Anal Chem 77:2176–2186. doi: 10.1021/ac048410j

121. Pringle SD, Giles K, Wildgoose JL, Williams JP, Slade SE, Thalassinos K, Bateman RH, Bowers MT, Scrivens JH (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. Int J Mass Spectrom 261:1–12. doi: 10.1016/j.ijms.2006.07.021

122. Gonzales GB, Raes K, Coelus S, Struijs K, Smagghe G, Van Camp J (2014) Ultra(high)-pressure liquid chromatography-electrospray ionization-time-of-flight-ion mobility-high definition mass spectrometry for the rapid identification and structural characterization of flavonoid glycosides from cauliflower waste. J Chromatogr A 1323:39–48. doi: 10.1016/j.chroma.2013.10.077

123. Stow SM, Causon TJ, Zheng X, Kurulugama RT, Mairinger T, May JC, Rennie EE, Baker ES, Smith RD, Mclean JA, Hann S, Fjeldsted JC (2017) An Interlaboratory Evaluation of Drift Tube Ion Mobility − Mass Spectrometry Collision Cross Section Measurements. Anal Chem 89:9048–9055. doi: 10.1021/acs.analchem.7b01729

124. Sun J, Baker A, Chen P (2011) Profiling the indole alkaloids in yohimbe bark with ultra-performance liquid chromatography coupled with ion mobility quadrupole time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 25:2591–2602. doi: 10.1002/rcm.5158

125. Yassin GH, Grun C, Koek JH, Assaf I, Kuhnert N (2014) Investigation of isomeric flavanol structures in black tea thearubigins using ultraperformance liquid chromatography coupled to hybrid quadrupole/ion mobility/time of flight mass spectrometry. J Mass Spectrom 1086–1095. doi: 10.1002/jms.3406

126. Stander MA, Van Wyk B, Taylor MJC, Long HS (2017) Analysis of Phenolic Compounds in Rooibos Tea (Aspalathus linearis) with a Comparison of

Flavonoid-Based Compounds in Natural Populations of Plants from Different Regions. J Agric Food Chem 65:10270–10281. doi: 10.1021/acs.jafc.7b03942

127. Jaiswal R, Jayasinghe L, Kuhnert N (2012) Identification and characterization of proanthocyanidins of 16 members of the Rhododendron genus (Ericaceae) by tandem LC-MS. J Mass Spectrom 47:502–515. doi: 10.1002/jms.2954

128. Chalet C, Hollebrands B, Janssen H, Augustijns P, Duchateau G (2018) Identification of phase-II metabolites of flavonoids by liquid chromatography – ion-mobility spectrometry – mass spectrometry. Anal Bioanal Chem 410:471–482

129. Xie C, Yu K, Zhong D, Yuan T, Ye F, Jarrell JA, Millar A, Chen X (2011) Investigation of Isomeric Transformations of Chlorogenic Acid in Buffers and Biological Matrixes by Ultraperformance Liquid Chromatography Coupled with Hybrid Quadrupole/Ion Mobility/Orthogonal Acceleration Time-of-Flight Mass Spectrometry. J Agric Food Chem 59:11078–11087

130. Kuhnert N, Yassin GH, Jaiswal R, Matei MF, Grün CH (2015) Differentiation of prototropic ions in regioisomeric caffeoyl quinic acids by electrospray ion mobility mass spectrometry. Rapid Commun Mass Spectrom 29:675–680

131. Warnke S, Seo J, Boschmans J, Sobott F, Scrivens JH, Bleiholder C, Bowers MT, Gewinner S, Schöllkopf W, Pagel K, Von Helden G (2015) Protomers of benzocaine: Solvent and permittivity dependence. J Am Chem Soc 137:4236–4242. doi: 10.1021/jacs.5b01338

132. Zhao B, Zhuang X, Pi Z, Liu S, Liu Z, Song F (2018) Determining the Effect of Catechins on SOD1 Conformation and Aggregation by Ion Mobility Mass Spectrometry Combined with Optical Spectroscopy. Am Soc Mass Spectrom 29:734–741. doi: 10.1007/s13361-017-1864-z

133. de Almeida NEC, Do TD, Lapointe NE, Tro M, Feinstein SC, Shea J, Bowers MT (2017) 1,2,3,4,6-penta-O-galloyl-B-d-glucopyranose binds to the N-terminal metal binding region to inhibit amyloid B-protein oligomer and fibril formation. Int J Mass Spectrom 420:24–34. doi: 10.1016/j.ijms.2016.09.018

134. De Almeida NEC, Do TD, Tro M, Lapointe NE, Feinstein SC, Shea JE, Bowers MT (2016) Opposing Effects of Cucurbit[7]uril and 1,2,3,4,6-Penta-O-galloyl-β-d-glucopyranose on Amyloid β25-35 Assembly. ACS Chem Neurosci 7:218–226. doi: 10.1021/acschemneuro.5b00280

135. Young LM, Cao P, Raleigh DP, Ashcroft AE, Radford SE (2014) Ion mobility spectrometry-mass spectrometry defines the oligomeric intermediates in amylin amyloid formation and the mode of action of inhibitors. J Am Chem Soc 136:660–670. doi: 10.1021/ja406831n

136. Zhang T, Zhang J, Derreumaux P, Mu Y (2013) Molecular Mechanism of the Inhibition of EGCG on the Alzheimer Aβ1–42 Dimer. J Phys Chem B 117:3993–4002. doi: 10.1021/jp312573y

137. Franceschi P, Vrhovsek U, Guella G (2011) Ion mobility mass spectrometric investigation of ellagitannins and their non-covalent aggregates. Rapid Commun Mass Spectrom 25:827–833

138. Groessl M, Azzollini A, Eugster PJ, Plet B, Wolfender J, Knochenmuss R (2014) Comparison of UHPLC-ESI-MS and Hadamard Transform Atmospheric Pressure Ion Mobility-ESI-MS for Rapid Profiling of Isomeric Flavonoids. Chimia (Aarau) 68:135–139. doi: 10.2533/chimia.2014.135

139. Krueger CA, Midey AJ, Kim T, Osgood M, Wu J, Wu C (2011) High performance ion mobility spectrometry as a fast and low cost green analytical technology part I : analysis of nutritional supplements. Int J IOn Mobil Spec 14:71–79. doi: 10.1007/s12127-011-0072-y

140. Midey AJ, Camacho A, Sampathkumaran J, Krueger CA, Osgood MA, Wu C (2013) High-performance ion mobility spectrometry with direct electrospray ionization (ESI-HPIMS) for the detection of additives and contaminants in food. Anal Chim Acta 804:197–206. doi: 10.1016/j.aca.2013.10.010

141. Hädener M, Kamrath MZ, Weinmann W, Groessl M (2018) High-Resolution Ion Mobility Spectrometry for Rapid Cannabis Potency Testing. Anal Chem 90:8764–8768. doi: 10.1021/acs.analchem.8b02180

142. Wang X, Zhou Q, Jiang D, Gong Y, Li E, Li H (2016) Ion mobility spectrometry as a simple and rapid method to measure the plasma propofol concentrations for intravenous anaesthesia monitoring. Sci Rep 6:1–7. doi: 10.1038/srep37525

143. Willems JL, Khamis MM, Mohammed W, Purves RW, Katselis G, Low NH, El-aneed A (2016) Analysis of a series of chlorogenic acid isomers using differential ion mobility and tandem mass spectrometry. Anal Chim Acta 933:164–174. doi: 10.1016/j.aca.2016.05.041

144. Troc A, Zimnicka M, Danikiewicz W (2015) Separation of catechin epimers by complexation using ion mobility mass spectrometry. J Mass Spectrom 50:542–548. doi: 10.1002/jms.3560

Declaration with signatures in possession of candidate and supervisor.

Declaration by the candidate:

With regard to Chapter 3, the nature and scope of my contribution were as follows:

Nature of contribution	**Extent of contribution (%)**
Performed the experiments, data analysis, co-wrote paper	60

The following co-authors have contributed to Chapter 3:

Name	**E-mail address**	**Nature of contribution**	**Extent of contribution (%)**
Tim Causon	tim.causon@boku.ac.at	Performed drift tube (DT) IM experiments, confirmed results on DT instrument; data analysis; editorial input	10
Harald Pasch	hpasch@sun.ac.za	Editorial input	5
André de Villiers	ajdevill@sun.ac.za	Data analysis, editorial input	25

Signature of candidate:

Date:24/1/2018

Declaration by co-authors:

The undersigned confirm that:

1. The declaration above accurately reflects the nature and extent of the contributions of the candidate and the co-authors to Chapter 3,
2. No other authors contributed to Chapter 3 besides those specified above, and
3. Potential conflicts of interest have been revealed to all interested parties and that the necessary changes have been made to use the material in Chapter 3 of this dissertation.

Signature	**Institutional affiliation**	**Date**
	BOKU Vienna	25/1/2018
	Stellenbosch University	25/1/2018
	Stellenbosch University	25/1/2018

Chapter 3

Comprehensive analysis of hydrolysable tannins by reversed phase and hydrophilic interaction chromatography coupled to ion mobility and high-resolution mass spectrometry, Part 1: Chestnut tannins

Abstract

In this two-part study, we report a methodology based on reversed phase LC (RP-LC) and hydrophilic interaction chromatography (HILIC) separations coupled to ion mobility (IM) and high-resolution mass spectrometry (HR-MS) for the detailed analysis of hydrolysable tannins. The current contribution presents the application of this approach to the analysis of an industrial chestnut (*Castanea sativa*, wood chips) tannin extract. A total of 38 molecular species, including a large number or isomers, were tentatively identified in this sample based on HR-MS and UV absorption spectral information as well as retention behaviour in both separation modes. In total, 128 and 90 isomeric species were resolved by RP- and HILIC-LC-IM-TOF-MS, respectively. The combination of low- and high collision energy mass spectral data with complementary chromatographic separations allowed tentative identification of twenty molecular species, comprising 78 isomers, in chestnut for the first time. Ion mobility resolved six new dimeric and trimeric vescalagin conformers with unique arrival (drift) times, including new conformers of roburin A-D which were not separated using either RP-LC or HILIC. HILIC was found to be the preferred separation mode for the analysis of vescalagin derivatives, while RP-LC is preferred for the analysis of ellagitannins with a cyclic glucose core. For the complete separation of the galloyl glucose species, comprehensive HILIC×RP-LC separation would be required.

3.1 Introduction

Hydrolysable tannins form part of the polyphenol family of vegetable tannins, which also include condensed tannins [1]. The term originates from the production of ellagic and/or gallic acid by hydrolysis of ester bonds of the corresponding tannins [2–4]. The main use for hydrolysable tannins is in the transformation of hides and skins into leather in a process called tanning [3]. The two most commonly used commercial tannin extracts for this purpose are chestnut and quebracho, with chestnut in particular being used for tanning of leather soles. Hydrolysable tannins also exhibit biological activities, including anti-cancer [5, 6] and anti-HIV properties [7, 8], and have been used in the treatment of diabetic complications [9].

Two main classes of hydrolysable tannins are distinguished: ellagitannins and gallotannins. Ellagitannins contain a hexahydroxydiphenoyl (HHDP) unit [10], which is formed by the oxidative C-C coupling between adjacent galloyl units esterified to a polyol (glucose) core (**Figure 3.1A**). Also esterified to the polyol core of some ellagitannins is a nonahydroxyterphenoyl (NHTP) group [11]. Hydrolysis of ellagitannins produces ellagic acid due to spontaneous lactonisation or intra-molecular esterification [12]. Ellagitannins can further be subdivided into structures containing ‘open’ or ‘closed glucose cores [5]. Gallotannins contain galloyl units which are esterified to a glucose core, without C-C coupling between adjacent galloyl units [13] (**Figure 3.1B**). All hydrolysable tannins originate from the biosynthetic intermediate, β-1,2,3,4,5-pentagalloyl-D-glucose, either as *in situ* metabolites of this intermediate [14–17] and/or as products formed during the extraction process of hydrolysable tannins [18]. Both ellagitannins and gallotannins comprise large numbers of isomeric structures. This isomeric diversity makes the characterisation of hydrolysable tannins extremely challenging, and many minor constituents remain unidentified.

Ultraviolet (UV) absorbance spectroscopy can be used to differentiate between ellagitannins and gallotannins based on the presence of an absorption band at ~280 nm in the case of the latter. Furthermore, differentiation between ‘mixed’ ellagitannins which contain both HHDP and galloyl units in different ratios is possible based on their relative absorbance at this wavelength [17, 19], while ellagic acid derivatives are distinguishable based on their absorbance at ~355 nm [2]. UV spectroscopy, however, provides limited information on the detailed molecular structures of hydrolysable

tannins, and since these compounds typically occur as complex mixtures in natural products, separation of the individual species is essential.

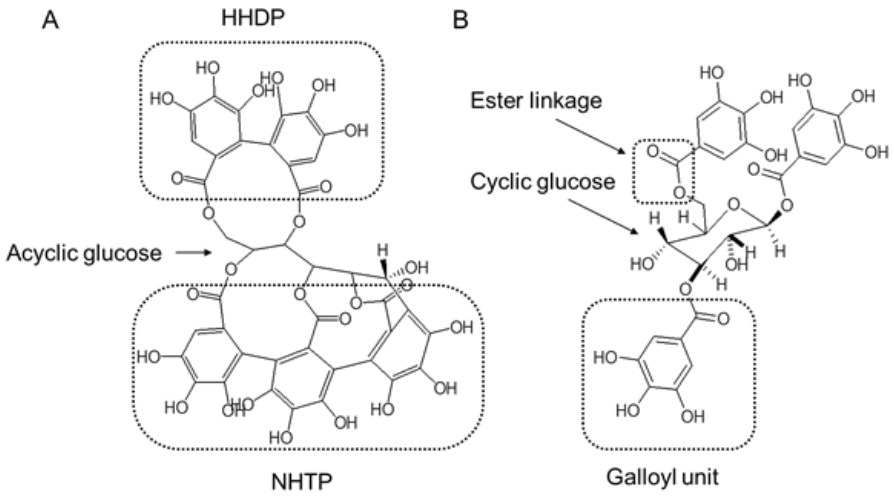


Figure 3.1: Representative structures of the two classes of hydrolysable tannins found in chestnut: (**A**) an ellagitannin (vescalagin, **17[a]**) and a (**B**) gallotannin (1,3,6-trigalloylglucose, **30**). HHDP: hexahydroxydiphenoyl, NHTP: nonahydroxytriphenoyl. For structures of the compounds identified in the current work, the reader is referred to **Figure S1**.

In terms of the chromatographic separation of hydrolysable tannins, early work has been reviewed by Mueller-Harvey [20]. Vescalagin (**17[a]**) and castalagin (**17[b]**), the two dominant ellagitannins in chestnut, were first isolated and characterised by Mayer *et al*. [21] (the structures have since been revised [22–24]). These authors also illustrated that hydrolysis of vescalagin and castalagin produces vescalin and castalin, respectively, as well as ellagic acid [2, 12]. The group of Okuda [13] performed important work on the analysis of hydrolysable tannins by utilising RP-LC [25], normal phase liquid chromatography (NP-LC) [26], centrifugal partitioning chromatography (CPC) and gel permeation chromatography (GPC) [27] for the fractionation of ellagitannins. These authors demonstrated that NP-LC was particularly useful for the separation of oligomeric tannins, whereas RP-LC allows the separation of anomeric ellagitannins such as vescalagin and castalagin [28]. Scalbert and co-workers [29] developed a method comprising size exclusion chromatography (SEC) followed by

preparative RP-LC to isolate and characterise ellagitannins, including vescalagin and castalagin. The same methodology was subsequently applied to identify grandinin (vescalagin-C-pentoside) and roburins A-E in oak wood [11, 30] and chestnut [31] by NMR. Viriot *et al.* [31] also detected trace levels of grandinin and roburins B, C and E in chestnut, and hypothesised that the dimerisation of vescalagin to produce roburin A is a non-enzymatic process and occurs with or without the addition of acid. More recently, RP-LC coupled to electrospray ionisation (ESI) MS [32] and tandem MS [33] have become popular methods for the analysis of hydrolysable tannins. For example, monogalloyl- to pentagalloyl-glucose gallotannins species were tentatively identified in chestnut by means of RP-LC-ESI-MS [32].

Stand-alone mass spectrometry has also been used for hydrolysable tannin analysis, including secondary ion mass spectrometry [34] and especially matrix-assisted laser desorption ionisation-time-of-flight (MALDI-TOF) MS, which has extensively been used in the analysis of hydrolysable tannins [35–38], including chestnut [18, 39, 40]. Pasch and Pizzi [18] reported that the *in situ* macromolecular structure of chestnut tannins could be composed of β-1,2,3,4,5-pentagalloyl glucose monomers, while Pizzi and co-workers [39] showed that the sugar moieties could be covalently bonded in chestnut.

More recently, ion mobility-mass spectrometry (IM-MS) has also found application for the analysis of hydrolysable tannins. Ion mobility entails the gas-phase separation of ions in an electric field according to size-to-charge ratio due to their collision with an inert buffer gas [41–44]. Franceschi and co-workers [45] reported the strong mutual interaction (hydrogen bonding) of ellagitannin oligomeric standards, resulting in multiply charged aggregates displaying different arrival times when measured using ‘travelling wave’ ion mobility (TWIMS) [46]. Since IM provides separation on the millisecond timescale, it is ideally compatible as an intermediate separation step between chromatography and MS [47, 48] which provides alternative selectivity to either of these techniques. To the best of our knowledge, IM has not to date been applied in combination with chromatographic separation to the analysis of hydrolysable tannins.

It is clear from the above that the analysis of hydrolysable tannins is a challenging task. For the complete structural characterisation, NMR is essential, but requires time-

consuming fractionation and purification procedures [30, 49]. Identification of minor constituents is not feasible in this manner. For this reason, there is interest in analytical methods suitable for the screening of hydrolysable tannins in natural products. The large number of potential isomers means that stand-alone mass spectrometry is limited, and LC-UV-MS methods are of increasing importance for this purpose [17]. However, conventional RP-LC methods used for hydrolysable tannin analyses also suffer from incomplete resolution of complex samples due to the large numbers of compounds involved and limited selectivity. To overcome these limitations, we report in this two-part study, a methodology based on the use of HILIC and RP-LC in combination with IM and high resolution MS for the analysis of hydrolysable tannins. In the current work the application of these methods to chestnut analysis is discussed. We also report for the first time the collisional cross sections ($^{TW}CCS_{N2}$) of the tentatively identified ellagitannins and gallotannins, which offer an additional means for compound identification.

3.2 Experimental

3.2.1 Materials and reagents

An industrial chestnut sample produced by extracting wood chips with 2% (m/v) $NaHSO_3$ and 0.5% $NaHCO_3$ in hot water (70-80°C) was investigated. HPLC grade acetonitrile and formic acid were purchased from Sigma-Aldrich (Johannesburg, South Africa) and deionised water was obtained using a Milli-Q water purification system (Millipore, Milford, MA). Poly-DL-alanine purchased from Sigma-Aldrich was used for IM calibration (Section 3.2.6).

3.2.2 Instrumentation

All analyses were performed on an Acquity UPLC system equipped with a binary pump, autosampler, column oven, photodiode array (PDA) detector (500 nL flow cell, 10 mm path length) connected via an ESI source to a Synapt-G2 quadrupole time-of-flight (Q-TOF) mass spectrometer (Waters, Milford, MA, USA). UV detection was performed from 230 to 500 nm at a 20 Hz acquisition rate. MS data were acquired in negative ionisation mode with a capillary voltage of −2.5 kV, a cone voltage of 15 V, a source temperature of 120°C and an extraction cone voltage of 4.0 V. Data were

acquired from 100 to 2000 amu with a scan time of 0.2 seconds (5 Hz). For LC–MS^E acquisition a collision energy ramp of 10–30 V was used. The source and desolvation temperatures were set to 120 and 275°C, respectively. Nitrogen was used as desolvation gas at a flow rate of 650 L/h. The instrument was calibrated using a sodium formate solution, and leucine enkephalin (m/z = 554.2615) was used as the lock mass calibrant.

For IM measurements the pressures in the helium and N_2 mobility cells were maintained using 180 and 90 mL/min flows, respectively. The mobility T-Wave was operated at a velocity of 448 m/s with a wave height of 37.1 V and a transfer velocity of 380 m/s.

3.2.3 Chromatographic conditions

For both RP-LC and HILIC separations a binary mobile phase of (A) water containing 0.1% (v/v) formic acid and (B) acetonitrile was used. Separations were performed at 30°C. The sample was dissolved at 20 mg/mL in 50% methanol (v/v), and 2 µL was injected in full loop injection mode with weak and strong needle washes comprising of acetonitrile and MeOH/H_2O (50/50, v/v) for both HILIC and RP-LC.

3.2.4 HILIC analyses

Separations were performed on an XBridge Amide (150 × 4.6 mm i.d., 2.5 µm) column (Waters). A linear gradient was performed as follows at a flow rate of 1 mL/min: 10–20% A (0.0–20.0 min), 20–50% A (20.0–30.0 min), 50% A (30.0–35.0 min). A 1:3 split was installed between the PDA and MS detectors.

3.2.5 RP-LC analyses

Separations were performed on a Kinetex C_{18} (100 × 2.1 mm i.d., 1.7 µm) superficially porous column (Phenomenex, Torrance, USA). A linear gradient was performed as follows: 1.0% B (0.0–2.5 min), 1.0–15.0% B (2.5–10.0 min), 15-60% B (10.0–14.0 min), 60% B until 20.0 minutes. The flow rate was 0.4 mL/min, with the total flow directed to the MS source.

3.2.6 Collisional cross section determination

Poly-DL-analine was prepared at a concentration of 10 ppm in deionised water, and wave height and velocity were optimised by direct injection into the MS. To determine $^{TW}CCS_{N2}$ (Ω) values from arrival times, the protocol outlined by Ruotolo *et al*. [50] was followed. Briefly, the recorded arrival (drift) time (t_d) was corrected for the mass to charge ratio (*m/z*). Ω values for poly-DL-alanine [51] were also corrected ($Ω'$) for the ion charge (z) as well as the reduced mass (μ). A calibration curve of ln t'_d vs. ln $Ω'$ was used to obtain the exponential factor, X, which was then used to obtain the double-corrected arrival time, t''_d, which was plotted against the known Ω values of poly-DL-aniline ions in the range 230-1011 m/z. This calibration curve was used to calculate Ω values of the identified hydrolysable tannins. To confirm the validity of the calibration approach, $^{TW}CCS_{N2}$ values for several standard phenolic compounds were determined and compared to literature values, and also to values determined for the same compounds on a drift tube ion mobility instrument (Agilent 6560 Ion Mobility Q-TOF).

3.2.7 Data processing

Data acquisition and processing were performed using MassLynx (v. 4.1) and Driftscope (v. 2.1) software (Waters). Driftscope was used to obtain mass spectra filtered as a function of arrival time for both low and high collision energy data.

3.2.8 RP-LC with low-field drift tube ion mobility-mass spectrometry measurements

RP-LC separation was performed using a Zorbax Eclipse Plus C_{18} Rapid Resolution column (2.1 × 50 mm, 1.8 μm) at a temperature of 40°C. The separation was performed using a binary mobile phase comprising (A) 0.1% (v/v) formic acid in water, and (B) methanol. Using a flow rate of 0.2 mL/min, an initial composition of 98% A was held for 0.17 minutes, followed by a gradient from 2-18% B in 0.17-7.2 minutes, then increasing to 25% from 7.2-12 minutes, and finally to 50% B from 12-13.8 minutes. A washing step was then performed by increasing to 90% B and holding for 1 minute. A re-equilibration time of 3 minutes was used (total run time of 21 minutes).

Nitrogen was used as drying gas at a temperature of 360 °C, a sheath gas temperature of 225 °C and a sheath gas flow rate of 13 L/min. The nebuliser gas pressure was 30 psi, the MS capillary voltage was -3500 V, the nozzle voltage -500 V and the

fragmentor was set to -275 V. Following tuning in the 2 GHz extended dynamic range mode with a mass range of 50-1700 m/z, mass calibration was undertaken immediately prior to measurements using the supplied tune mixture of the manufacturer.

The IM trapping funnel was operated with an accumulation time of 10 000 µs and released packages of ions every 60 ms with a trap release time of 150 µs set within the software. The drift tube was operated with an absolute entrance voltage of -1574 V and an exit voltage of -224 V with a drift tube pressure of 3.95 Torr and a temperature of 26-28 °C (ambient) using high purity nitrogen (99.999%) as the collision gas. The acquisition settings were adjusted to yield 30 ion mobility transients per frame corresponding to 0.5 ion mobility frames per second.

A single-field calibration approach was employed using $^{DT}CCS_{N2}$ values and methodology established in a recent study [52].

3.3 Results and discussion

3.3.1 Chromatographic separation

In the first part of the study, the chromatographic separation of the chestnut tannin sample was investigated. RP-LC and HILIC separations were used and compared in terms of the number of resolved peaks and selectivity. Base peak ion (BPI) chromatograms obtained for the RP-LC- and HILIC-IM-MS analyses of the chestnut tannin extract are presented in **Figure 3.2** (compound numbers in this figure correspond to **Tables 3.1-3.3**). While different column dimensions, flow rates and gradient times were used for HILIC and RP-LC separations, good separation performance was obtained for both. For RP-LC, a superficially porous phase was used, whereas in HILIC an amide phase was used due to the known good performance of both these phases for phenolic compounds [53, 54]. To our knowledge, this is the first application of HILIC to the analysis of hydrolysable tannins. From **Figure 3.2**, the complementary selectivities of HILIC and RP-LC separations are clearly evident from the respective elution orders observed for the identified compounds. This proved highly beneficial in the current work, as compounds co-eluting in either separation

were often resolved in the other. The compounds tentatively identified are summarised in **Tables 3.1**-**3.3**. Detailed discussions on the step-by-step identification of different classes of hydrolysable tannins detected in chestnut are presented in the following sections.

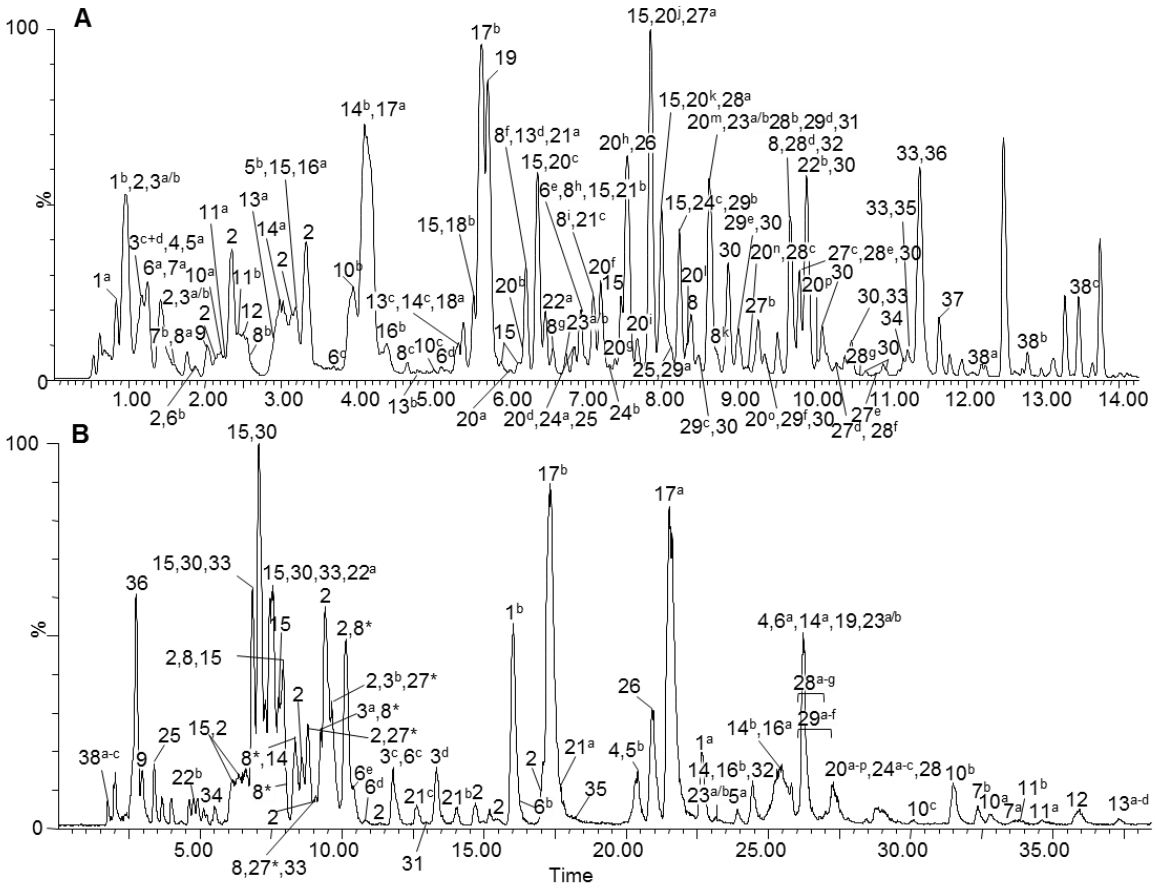


Figure 3.2: Base peak ion chromatograms obtained for the (**A**) RP-LC- and (**B**) HILIC-IM-Q-TOF-MS analyses of chestnut hydrolysable tannins. Peak labels correspond to **Tables 3.1-3.3**. For experimental conditions, refer to **Section 3.2.3**.*Indicates isomers which could not be matched between RP-LC and HILIC.

3.3.2 Identification of chestnut hydrolysable tannins

3.3.2.1 Ellagitannins with an open glucose core

Due to the lack of reference compounds, tentative identification of hydrolysable tannins in chestnut was based on comparison of elution orders in RP-LC and NP-LC, UV spectral information, and especially high resolution low- and high collision energy

MS data compared to previous literature reports [19, 29–32, 45, 55–59]. Although MS/MS experiments were not performed in the current study, the use of IM allows the option of obtaining 'clean' mass spectra by filtering data for (partially) co-eluting compounds in terms of arrival times; unless otherwise stated, all mass spectra reported in this contribution have been filtered in this manner. **Table 3.1** and **Figure 3.3** summarise the vescalagin and castalagin derivatives tentatively derivatives identified in chestnut in this manner.

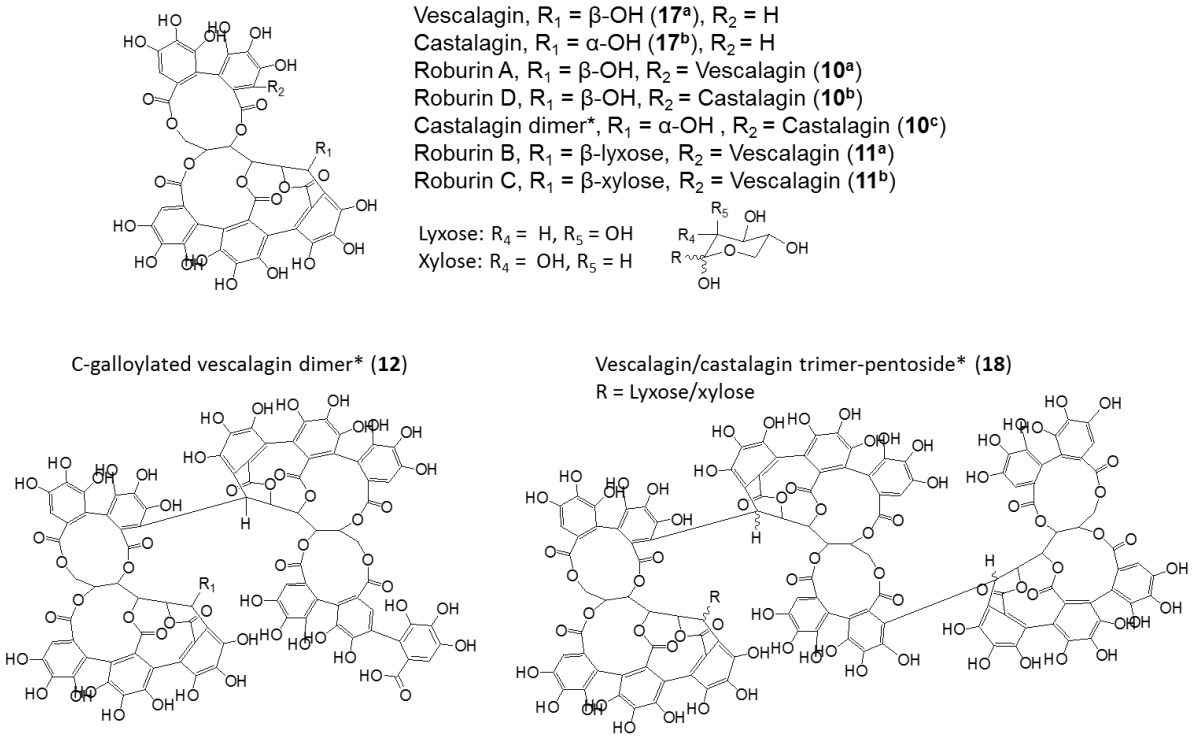


Figure 3.3: Vescalagin/castalagin derivatives identified in chestnut by HILIC and RP-LC-IM-MS. Compounds **10, 11, 12,** and **18** displayed multiple arrival times. *Indicates compounds tentatively identified for the first time.

Collisional induced dissociation (CID) spectra obtained in MS^E mode can be used to differentiate between isomeric compounds such as vescalagin (**17a**) and castalagin (**17b**), which only differ at the C-1 anomeric carbon of the glucose moiety [32, 56]. MS^E spectra for both isomers are presented in **Figure 3.4**, and a general fragmentation scheme for these compounds is presented in **Figure S2** in the SI. Vescalagin, with its

hydroxyl group in the β position (S configuration), produces a fragment ion at 915 m/z $[M-H_2O-H]^{1-}$ which is present in much lower abundance in the MS^E spectrum of castalagin [17]. In contrast, fragmentation of castalagin predominately involves the loss of the HDDP unit, producing ions at 631 m/z $[M-HHDP-H]^{1-}$ and 301 m/z $[HHDP-H]^{1-}$. The fragment ions at m/z 915 and 631 were, therefore, used to identify the known ellagitannins, as well as to tentatively identify new vescalagin and castalagin derivatives presented in **Table 3.1**. Both the molecular ion (933 m/z $[M-H]^{1-}$) and the $[M-HHDP-H]^{1-}$ fragment (631 m/z) undergo losses of water and decarboxylation (**Figure 3.4A**).

During the hydrolysis of hydrolysable tannins containing an HHDP unit, ellagic acid is formed, followed by spontaneous lactonisation. The same transformation also occurs in the gas phase during the fragmentation of ellagitannins containing a HHDP unit [57] (**Figure S2**). Fragmentation of the HHDP unit produces diagnostic fragment ions at 301 $[ellagic\ acid-H]^{1-}$, 275 $[ellagic\ acid-H-CO]^{1-}$ and 249 $[ellagic\ acid-H-2CO]^{1-}$ m/z, which may be used to identify ellagitannins containing the HHDP functional group.

It can furthermore be observed from **Figure 3.4** that both compounds undergo more extensive fragmentation under HILIC conditions when using identical MS parameters. The same behaviour was noted for all compounds identified in the current work (illustrated in **Figure S3** for tellimagrandin I (**27**) as example of a closed glucose core structure). This could be due to more extensive in-source fragmentation under HILIC conditions as a result of improved desolvation from the organic-rich HILIC mobile phase, although this aspect was not investigated further.

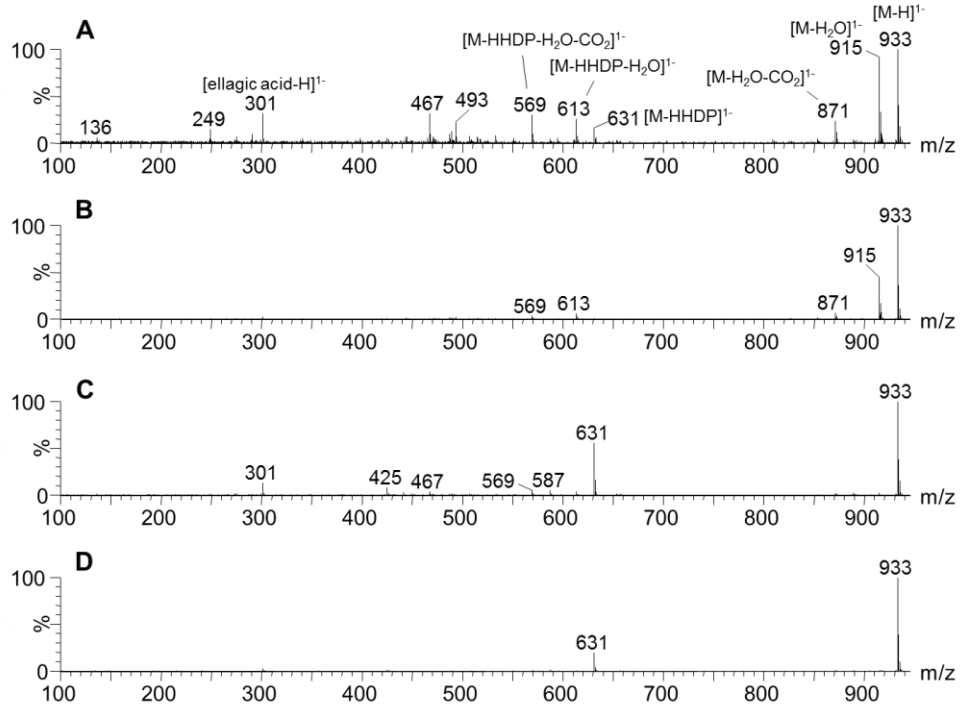


Figure 3.4: MS^E spectra obtained for (**A** and **B**) vescalagin (**17[a]**) and (**C** and **D**) castalagin (**17[b]**) under HILIC (**A** and **C**) and RP-LC (**B** and **D**) conditions.

Table 3.1: Castalagin and vescalagin derivatives tentatively identified in chestnut using RP-LC- and HILIC-IM-HR-MS.

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
Vescalin (1[a])	$C_{27}H_{20}O_{18}$	0.83	22.65	230	631.0582	1.74	3.80	219	nd*				613.0837, 587.1044, 569.0948, 471.0238, 467.0552, 449.0458, 425.0387, 327.0351, 301.0193, 273.0213, 119.9524
Castalin (1[b])	$C_{27}H_{20}O_{18}$	0.97	16.03	230	631.0591	3.10							613.0837, 587.1056, 569.0928, 471.0255, 467.0557, 425.0407, 327.0363, 301.0268, 288.0444, 273.0180, 193.0470
Di-hydrated Vescalagin (5[a])	$C_{41}H_{30}O_{28}$	1.19	23.95	230	969.0884	4.02	6.35	283	484.0412	5.89	2.00	312	667.0807, 300.9994, 275.0194
Castalagin-castalin/vescalin dimer (7[a])	$C_{68}H_{44}O_{43}$	1.42	33.87	235	nd		nd		773.0576	3.36	3.17	393	933.0696,631.0585, 613.0507, 569.0596, 425.0140, 399.0348, 300.9993, 275.0201
Castalagin-castalin/vescalin dimer (7[b])	$C_{68}H_{44}O_{43}$	1.55	32.36						773.0566	2.07	3.17	393	
Roburin A (10[a])	$C_{82}H_{50}O_{51}$	2.21	32.82	232	nd		nd		924.0615	3.68	3.52 4.35	414 462	933.0660, 915.0574, 871.0725, 613.0455, 569.0559, 493.0059, 300.9978, 275.0195, 249.0402
Roburin B (11[a])	$C_{87}H_{58}O_{55}$	2.26	34.63	232	nd		nd		990.0798	0.51	3.93 4.69	438 482	1065.1102, 915.0570, 645.0695, 613.0456, 569.0576, 300.9978, 275.0195
Roburin C (11[b])	$C_{87}H_{58}O_{55}$	2.43	33.70	233	nd		nd		990.0820	2.73	3.93 4.69	438 480	1065.1134, 915.0563, 645.1835, 613.0479, 569.0631, 300.9999, 275.0176
C-galloylated vescalagin dimer (12)	$C_{89}H_{54}O_{55}$	2.48	35.93	234	nd		nd		1000.0666	3.00	4.00 4.76	442 484	1085.0787, 1065.1134, 1041.0891, 915.0566, 871.0688, 613.0479, 331.0669, 300.9987, 275.0178, 271.0458
Vescalagin/castalagin-pedunculagin dimer (13[a])	$C_{75}H_{48}O_{47}$	2.93	37.32	230	nd		nd		849.0600	-0.59	3.66	423	co-elution*

Table 3.1 (continued).

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
Grandinin (14[a])	$C_{46}H_{34}O_{30}$	2.98	26.24	nd	1065.1099	3.97	7.59	310	532.0496	1.32	2.28	333	975.0757, 510.0554^{2-}, 487.0321^{2-}, 457.0320^{2-}, 300.9981, 275.0185, 249.0393, 169.0127
Di-hydrated Castalagin (5[b])	$C_{41}H_{30}O_{28}$	3.16	20.4	230	969.0892	4.85	6.35	283	484.0387	0.72	2.00	312	667.0816, 300.9981, 275.0191
Vescavaloninic acid (16[a])	$C_{48}H_{30}O_{31}$	3.25	25.57	nd	1101.0718	2.27	8.00	318	550.0328	3.64	2.35	338	co-elution
Roburin D (10[b])	$C_{82}H_{50}O_{51}$	3.95	31.50	232	nd		nd		924.0612	3.35	3.52 4.35	414 464	933.0651, 915.0548, 871.0647,631.0598 613.0459, 493.0043, 300.9976
Roburin E (14[b])	$C_{46}H_{34}O_{30}$	4.10	25.39	nd	1065.1072	1.43	7.59	310	532.0488	-0.19	2.28	333	975.0800, 487.0304^{2-}, 457.0232^{2-}, 300.9977, 275.0198, 249.0385, 169.0132
Vescalagin (17[a])	$C_{41}H_{26}O_{26}$	4.12	21.51	234	933.0659	2.67	6.62	290	466.0288	2.15	1.93	306	915.0549, 871.0648, 613.0468, 569.0730, 493.0043, 467.0254, 457.0234^{2-}, 300.9977, 275.0198, 249.0386, 231.0292, 169.0134, 155.0341, 137.0244, 125.0240
Castavaloninic acid (16[b])	$C_{48}H_{30}O_{32}$	4.38	24.41	nd	1101.0692	-0.09	8.00	318	550.0311	0.55	2.35	338	1057.0760, 425.0185
Vescalagin/castalagin-pedunculagin dimer (13[b])	$C_{75}H_{48}O_{47}$	4.83	37.32	230	nd		nd		849.0610	0.59	3.66	423	co-elution
Vescalagin/castalagin dimer (10[c])	$C_{82}H_{50}O_{51}$	5.01	30.01	nd	nd		nd		924.0577	-0.43	4.00	442	915.0557, 613.0464, 300.9986
Vescalagin/castalagin-pedunculagin dimer (13[c])	$C_{75}H_{48}O_{47}$	5.26	37.32	230	nd		nd		849.0600	-0.59	3.66	423	co-elution
Vescalagin/castalagin trimer-pentoside (18[a])	$C_{128}H_{82}O_{80}$	5.30	42.40	nd	nd		nd		1448.1034	-4.70	6.35 7.11	560 593	1831.0481, 1448.6106, 1382.5907, 1065.1064, 933.0693, 915.0538, 753.0581, 631.1832
Vescalagin/castalagin-pentoside (14[c])	$C_{46}H_{34}O_{30}$	5.31	nd	nd	1065.1046	-1.01	7.59	310	532.0491	0.38	2.28	333	low sensitivity*

Table 3.1 (continued).

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
Vescalagin/castalagin trimer-pentoside (18[b])	$C_{128}H_{82}O_{80}$	5.57	38.94	nd	nd		nd		1448.1047	-3.80	6.35 7.11	560 593	co-elution
Castalagin (17[b])	$C_{41}H_{26}O_{26}$	5.63	17.28	233	933.0651	1.82	6.62	290	466.0290	2.57	1.93	306	631.0585, 587.0677, 569.0602, 425.0127, 457.0222^{2-},300.9980 275.0188, 249.0392, 245.0085, 231.0283, 169.0128, 125.0231
O-galloylated castalagin/vescalagin (19)	$C_{48}H_{30}O_{30}$	5.72	26.22	nd	1085.0771	2.51	7.80	314	542.0338	0.92	2.35	338	1041.0895, 915.0546, 520.0392^{2}, 493.0039, 467.0261, 457.0212^{2-}, 300.9978, 275.0188, 249.0391, 169.0122, 125.0233
Vescalagin/castalagin- pentoside (14[d])	$C_{46}H_{34}O_{30}$	5.94	nd	nd	1065.1047	-0.92	7.59	310	532.0488	-0.19	2.28	333	low intensity
Vescalagin/castalagin-digalloyl glucose dimer (20[a])	$C_{61}H_{44}O_{39}$	5.96	27.30	274	nd		9.94	355	699.0674	3.15	2.97	381	co-elution
Vescalagin/castalagin-digalloyl glucose dimer (20[b])	$C_{61}H_{44}O_{39}$	6.16	27.30	274	nd		9.94	355	699.0671	2.72	2.97	381	co-elution
Vescalene (21[a])	$C_{27}H_{18}O_{17}$	6.21	17.43	251, 374	613.0475	1.47	3.66	215					523.0154, 465.0127, 300.9982
Vescalagin/castalagin-pedunculagin dimer (13[d])	$C_{75}H_{48}O_{47}$	6.24	37.32	230	nd		nd		849.0651	5.42	3.66	423	low sensitivity
Vescalagin/castalagin-digalloyl glucose dimer (20[c])	$C_{61}H_{44}O_{39}$	6.34	27.30	274	1399.1415	2.43	9.94	355	699.0682	4.29	2.97	381	co-elution
Valoneic acid dilactone (22[a])	$C_{21}H_{10}O_{13}$	6.47	7.52	254, 374	469.0053	2.13	2.90	192					469.0064, 425.0161, 407.0024, 301.0000, 169.0137
Vescalagin/castalagin-digalloyl glucose dimer (20[d])	$C_{61}H_{44}O_{39}$	6.85	27.30	274	1399.1362	-1.36	9.94	355	699.0652	0.00	2.97	381	co-elution
Vescalagin/castalagin-tellimagrandin I dimer (24[a])	$C_{75}H_{50}O_{47}$	6.86	27.25	nd	nd		nd		850.0692	1.06	3.52	415	915.0563, 785.0776, 300.9980

Table 3.1 (continued).

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
Vescalene (21[b])	$C_{27}H_{18}O_{17}$	6.99	14.02	251, 375	613.0491	4.11	3.66	215	nd				300.9992
Vescalagin/castalagin-digalloyl glucose dimer (20[e])	$C_{61}H_{44}O_{39}$	7.05	27.30	274	1399.1394	0.93	9.94	355	699.0672	2.86	2.97	381	co-elution
Vescalene (21[c])	$C_{27}H_{18}O_{17}$	7.10	12.64	251, 376	613.0485	3.13	3.66	215	nd				300.9996
Vescalagin/castalagin-digalloyl glucose dimer (20[f])	$C_{61}H_{44}O_{40}$	7.20	27.30	274	1399.1382	0.07	9.94	355	699.0671	2.72	2.97	381	915.0588, 483.0790, 331.0665, 313.0566, 300.9994, 271.0453, 249.0396, 211.0254, 193.0141
Vescalagin/castalagin-tellimagrandin I dimer (24[b])	$C_{75}H_{50}O_{47}$	7.31	27.25	nd	nd		nd		850.0720	4.35	3.52	415	915.0575, 785.0829, 300.9980
Vescalagin/castalagin-digalloyl glucose dimer (20[g])	$C_{61}H_{44}O_{39}$	7.42	27.30	274	nd		9.94	355	699.0676	3.43	2.97	381	co-elution
1-O-trihydroxybenzene castalagin (26)	$C_{47}H_{30}O_{28}$	7.55	20.92	nd	1041.0892	3.94	7.45	307	520.0412	4.42	2.28	333	915.0579, 739.0784, 613.0505, 483.0779, 493.0072, 467.0280, 457.0242^{2-}, 300.9987, 275.0194, 249.0397, 169.0138, 125.0231
Vescalagin/castalagin-digalloyl glucose dimer (20[h])	$C_{61}H_{44}O_{39}$	7.61	27.30	274	1399.1357	-1.72	9.94	355	699.0690	5.44	2.97	381	co-elution
Vescalagin/castalagin-digalloyl glucose dimer (20[i])	$C_{61}H_{44}O_{39}$	7.68	27.30	274	1399.1360	-1.50	9.94	355	699.0674	3.15	2.97	381	co-elution
Vescalagin/castalagin-digalloyl glucose dimer (20[j])	$C_{61}H_{44}O_{39}$	7.78	274		nd		9.94	355	699.0676	3.43	2.97	381	co-elution
Vescalagin/castalagin-digalloyl glucose dimer (20[k])	$C_{61}H_{44}O_{39}$	8.01	274		1399.1272	-7.79	9.94	355	699.0653	0.14	2.97	381	co-elution
Vescalagin/castalagin-tetragalloyl glucose dimer (28[a])	$C_{75}H_{52}O_{47}$	8.05	26.11/26.44/26.68[#]	nd	nd		nd		851.0771	1.17	3.45	410	co-elution
Vescalagin/castalagin-trigalloyl glucose dimer (29[a])	$C_{68}H_{48}O_{43}$	8.08	26.47	272	1551.1361	-8.38	10.42	363	775.0721	1.81	3.17	393	co-elution

Table 3.1 (continued).

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
Vescalagin/castalagin-trigalloyl glucose dimer (29[b])	$C_{68}H_{48}O_{43}$	8.20	26.47	272	1551.1361	-8.38	10.42	363	775.0704	-0.39	3.17	393	co-elution
Vescalagin/castalagin-tellimagrandin I dimer (24[c])	$C_{75}H_{50}O_{47}$	8.25	27.25	nd	nd		nd		850.0687	0.47	3.52	415	low intensity
Vescalagin/castalagin-digalloyl glucose dimer (20[l])	$C_{61}H_{44}O_{39}$	8.31	27.30	274	nd		9.94	355	699.0664	1.72	2.97	381	co-elution
Vescalagin/castalagin-trigalloyl glucose dimer (29[c])	$C_{68}H_{48}O_{43}$	8.49	26.47	272	1551.1533	2.71	10.42	363	775.0704	-0.39	3.17	393	co-elution
Vescalagin/castalagin-digalloyl glucose dimer (20[m])	$C_{61}H_{44}O_{39}$	8.58	27.30	274	nd		9.94	355	699.0658	0.86	2.97	381	co-elution
Vescalagin/castalagin-trigalloyl glucose dimer (29[d])	$C_{68}H_{48}O_{43}$	8.65	26.47	272	1551.1533	2.71	10.42	363	775.0711	0.52	3.17	393	co-elution
Vescalagin/castalagin-tetragalloyl glucose dimer (28[b])	$C_{75}H_{52}O_{47}$	8.65	26.11/26.44/26.68[#]	nd	nd		nd		851.0739	-2.58	3.45	410	co-elution
Vescalagin/castalagin-trigalloyl glucose dimer (29[e])	$C_{68}H_{48}O_{43}$	9.00	26.47	272	1551.1533	2.71	10.42	363	775.0731	3.10	3.17	393	1067.0593, 915.0527, 835.0693, 635.0906, 493.0075, 483.0780, 465.0656, 313.0528, 300.9956, 169.0124
Vescalagin/castalagin-digalloyl glucose dimer (20[n])	$C_{61}H_{44}O_{39}$	9.09	27.30	274	nd		9.94	355	699.0637	-2.15	2.97	381	co-elution
Vescalagin/castalagin-tetragalloyl glucose dimer (28[c])	$C_{75}H_{52}O_{47}$	9.16	26.11/26.44/26.68[#]	nd	nd		nd		851.0728	-3.88	3.45	410	915.0536, 787.1168, 300.9977
Vescalagin/castalagin-trigalloyl glucose dimer (29[f])	$C_{68}H_{48}O_{43}$	9.34	26.47	272	nd		10.42	363	775.0706	-0.13	3.17	393	1067.0773, 915.0519, 635.0902, 492.9997, 483.0787, 465.0663, 313.0566, 300.9979, 169.0119
Vescalagin/castalagin-digalloyl glucose dimer (20[o])	$C_{61}H_{44}O_{39}$	9.35	27.30	274	nd		9.94	355	699.0731	11.30	2.97	381	co-elution

Table 3.1 (continued).

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
Vescalagin/castalagin-tetragalloyl glucose dimer (28[d])	$C_{75}H_{52}O_{47}$	9.64	26.11/26.44/26.68[#]	nd	nd		nd		851.0798	4.35	3.45	410	co-elution
Vescalagin/castalagin-ellagic acid dimer (32)	$C_{55}H_{30}O_{33}$	9.67	24.45	250, 374	1217.0583	-0.66	8.35	325	608.0251	-0.99	2.62	357	915.0555, 493.0030, 467.0222, 300.9986, 275.0182, 249.0387, 169.0135
Vescalagin/castalagin-tetragalloyl glucose dimer (28[e])	$C_{75}H_{52}O_{47}$	9.85	26.11/26.44/26.68[#]	nd	nd		nd		851.0766	0.59	3.45	410	co-elution
Valoneic acid dilactone (22[b])	$C_{21}H_{10}O_{13}$	9.91	4.78	254, 375	469.0044	0.21	2.90	192	nd				469.0031, 425.0141, 299.9890, 169.0127
Vescalagin/castalagin-digalloyl glucose dimer (20[p])	$C_{61}H_{44}O_{39}$	9.97	27.30	274	nd		9.94	355	699.0863	nd	2.97	381	co-elution
Vescalagin/castalagin-tetragalloyl glucose dimer (28[f])	$C_{75}H_{52}O_{47}$	10.26	26.11/26.44/26.68[#]	nd	nd		nd		851.080	4.58	3.45	410	915.0862, 787.1168
Vescalagin/castalagin-tetragalloyl glucose dimer (28[g])	$C_{75}H_{52}O_{47}$	10.59	26.11/26.44/26.68[#]	nd	nd		nd		851.0781	2.35	3.45	410	915.0862, 787.1168
Ellagic acid-pentoside (34)	$C_{19}H_{14}O_{12}$	11.18	5.27	nd	433.0418	2.54	2.83	190					443.0423, 300.9981
Ellagic acid (36)	$C_{14}H_6O_8$	11.39	2.73	253, 367	300.9988	1.33	1.66	147					283.9965, 275.0202, 249.0396, 245.0101, 229.0112
Trimethoxy-ellagic acid (38[a])	$C_{17}H_{12}O_8$	12.19	1.76	247, 373	343.0447	-2.04	2.21	169					328.0287, 313.0052, 297.9824
Trimethoxy-ellagic acid (38[b])	$C_{17}H_{12}O_8$	12.79	1.76	247, 373	343.0462	2.33	2.21	169					328.0273, 313.0052, 297.9807
Trimethoxy-ellagic acid (38[c])	$C_{17}H_{12}O_8$	13.47	1.76	247, 373	343.0448	-1.75	2.21	169					328.0298, 313.0060, 297.9789

*t_d: arrival time; nd: not detected; co-elution/low intensity: no clear fragments obtained due to co-elution or low signal.

[#]Assignment of HILIC retention time was not possible due to fewer isomers detected in HILIC than RP-LC.

A useful feature of ion mobility is that product ions of singly and multiply charged species of a given molecule, which show multiple arrival times, can be compared by filtering high collision energy MS^E spectra based on arrival time. **Figure S4** illustrates this using vescalagin as an example (the same observations apply to several other compounds, including **17^b**, **14^a**, **14^b**, **26**, **16^{a+b}**). The MS^E spectra in **Figure S4** illustrate that the fragmentation spectrum obtained from the chromatographic trace (**A**) represents the sum of fragment ions derived from both the singly $[M-H]^{1-}$ (**B**) and doubly $[M-2H]^{2-}$ charged (**C**) ions. Fragmentation of the singly charged vescalagin molecular ion (933 m/z, $[M-H]^{1-}$) with an arrival time at 6.62 ms results in the high mass fragments at 915 ($[M-H_2O-H]^{1-}$), 871 ($[M-H_2O-CO_2-H]^{1-}$) and 613 m/z ($[M-HDDP-H_2O-H]^{1-}$). The doubly charged molecular ion at 466 m/z $[M-2H]^{2-}$ and an arrival time at 1.93 ms produces the fragment at 457 m/z ($[M-H_2O-2H]^{2-}$) in almost twice the abundance as the doubly charged molecular ion. Furthermore, lower mass fragments at 301, 275 and 249 m/z, corresponding to the diagnostic singly charged ellagic acid fragments, are also observed in the MS^E spectrum of the doubly charged species (**Figure S4C**). These data indicate that the extent of fragmentation is more severe for multiply charged ions. Furthermore, since more extensive fragmentation occurs under HILIC conditions, the MS^E spectrum of the doubly charged vescalagin ion (**Figure S4D**) shows the absence of the doubly charged molecular ion at 466 m/z ($[M-2H]^2$) as well as the dehydrated ion 475 m/z ($[M-H_2O-2H]^{2-}$), with the [ellagic acid-H]$^{1-}$ ion being the base peak in the MS^E spectrum.

Combining the observations made above regarding fragmentation spectra of singly (**Figure 3.4**) and doubly charged (**Figure S4**) species, castalagin and vescalagin can also be distinguished based on the high-energy mass spectra of their doubly charged species (**Figure S5**). These differ in the intensity of the 457 m/z ($[M-H_2O-2H]^{2-}$) ion, since the extent of water loss is significantly higher for vescalagin (**17^a**). Note that the arrival times (and $^{TW}CCS_{N2}$ values) for vescalagin and castalagin are identical (**Table 3.1**).

The utility of ion mobility for hydrolysable tannins becomes especially apparent when examining compounds **10**, **11**, **12** and **18** (**Table 3.1**), all of which are oligomeric vescalagin and/or castalagin derivatives for which multiple arrival times were observed. **Figure 3.5** shows the two-dimensional separation space obtained by the

RP-LC×IM and HILIC×IM separations of these oligomeric derivatives. Four of these compounds correspond to roburins A (vescalagin dimer, **10^{a}**), B (vescalagin-C-lyxoside-vescalagin dimer, **11^{a}**), C, (vescalagin-C-xyloside-vescalagin dimer, **11^{b}**) and D (vescalagin-castalagin dimer, **10^{b}**), first identified in chestnut by Viriot and co-workers [31]. These compounds were assigned (**Table 3.1**) based on their known elution order relative to vescalagin and castalagin in RP-LC [29–31]. Note that roburin A largely co-elutes with roburin B in RP-LC, whereas all four roburins are well resolved in HILIC. However, as is evident from **Figure 3.5**, two distinct arrival times were measured for each of the four roburin species, with identical arrival times in HILIC and RP-LC (similar behaviour was observed for several additional oligomeric ellagitannins, see further). For the isomeric pair roburin A and D, two arrival times were observed at 3.52 and 4.35 ms, whereas for the roburin B and C isomers, the arrival times were 3.93 and 4.69 ms, respectively. Considering that the two species of each roburin separated by ion mobility showed exactly the same retention time in RP-LC and HILIC (although the elution orders differed between these modes), it is unlikely that these are new isomeric compounds. A more likely explanation is that the two species detected for each of the roburins correspond to different conformers which are separated on the IM time scale.

It is unclear whether these species correspond to rotational conformers which are kinetically stable on the IM time-scale, or prototropic isomers formed as a result of multiple de-protonation sites [60]. To identify the structures of the relevant conformers, detailed molecular modelling would be required [61, 62]. Nevertheless, this is to our knowledge the first report of the separation of ellagitannin conformers, which was only possible by using the combination of two distinct separation methods in combination with IM. Indeed, IM proved especially valuable in this context, since the technique allowed assignment of peaks with identical arrival times in both separation modes. MS^{E} spectra of the conformers of roburin A and D with the shorter arrival times (3.52 ms, **Figure S6A** and **B**) confirm the presence of a castalagin unit in roburin D, identified by the 631 m/z ion [11, 30]. There was no observable difference in the fragmentation behaviour for the conformers for both roburins A and D at arrival times 3.52 and 4.35 ms.

One additional isomer of the roburin A/D type was detected at higher RP-LC and lower HILIC retention compared to roburin A and D (**10^{a+b}**). **Figure S6C** shows the fragmentation spectrum of this isomer (**10^{c}**, doubly charged molecular ion at 924 m/z), with a arrival time of 4.00 ms. Taking the elution order of vescalagin and castalagin as well as their related dimeric species into account, we propose that this could be a castalagin dimer (castalagin is more retained in RP-LC and less retained in HILIC than vescalagin). This hypothesis is supported by the relative intensity of the m/z 915 ion for **10^{c}**. Somewhat surprisingly, only one arrival time was detected for this compound, although this might be due to its low abundance. Further deductions on the structure of this compound could not be made based on the available data.

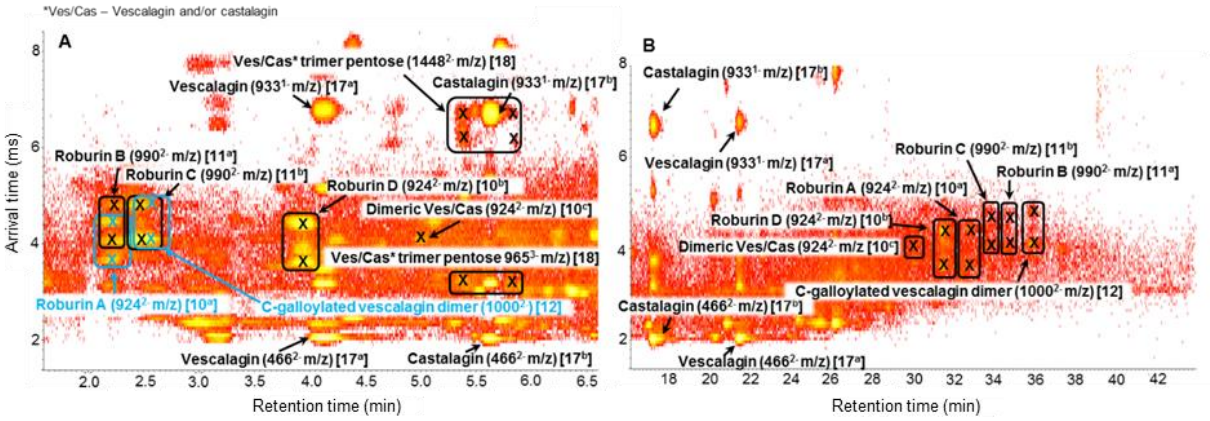


Figure 3.5: RP-LC×ion mobility (**A**) and HILIC×ion mobility (**B**) contour plots obtained for the separation of the monomeric and oligomeric vescalagin (Ves) and castalagin (Cas) derivatives. Compound numbers in square brackets [] correspond to **Table 3.1**.

The fragmentation spectra of the glycosylated isomers roburin B (**11^{a}**) and C (**11^{b}**) were similar, except for an additional ion at 343 m/z detected for roburin B (**Figure S7**). The ion at 1065 m/z [M-vescalagin-H]$^{1-}$ represents the glycosylated vescalagin fragment (identical to the molecular ion of grandinin or roburin E, **14**), and the 915 m/z [M-grandinin-H]$^{1-}$ and 613 m/z [M-grandinin-ellagic acid-H]$^{-}$ ions correspond to the vescalagin fragments of roburin B and C (the fragmentation of these compounds is illustrated in **Figure S8**).

A galloylated vescalagin dimer (**12**) is a newly proposed species detected as a doubly charged ion at 1000 m/z $[M-2H]^{2-}$, displaying multiple arrival times at 4.00 and 4.76 ms. Both these arrival times are virtually identical to those measured for roburin B and C, which is not unexpected considering the similarities between the structures of these compounds. Furthermore, **12** partially co-elutes with roburin C in RP-LC (*cf.* **Figure 3.5A**), which might explain why this compound has not been reported previously. Fragmentation of **12** produces ions at 915, 871 and 613 *m/z*, which indicates that both units comprise vescalagin (because of the absence of the 631 m/z ion, **Figure 3.6**). (Fragmentation behaviour was identical for both arrival times, suggesting that these species correspond to different conformers of the same molecule). The product ion at 1085 *m/z* representing the galloylated vescalagin moiety $[M\text{-vescalagin}]^{1-}$ points to the vescalagin monomeric units being coupled via a C-O-C bond as opposed to the much more common oxidative C-C coupling (**Figure 3.6**). This ion also undergoes a loss 44 m/z (CO_2, 1041 m/z), which indicates the presence of a free carboxylic acid group. In contrast to the valoneoyl group observed in vescavaloninic and castavaloninic acids (**16**, **Figure S1**) [17], however, the galloyl group is likely attached to an HHDP hydroxyl via an *o*- or *p*-hydroxyl group.

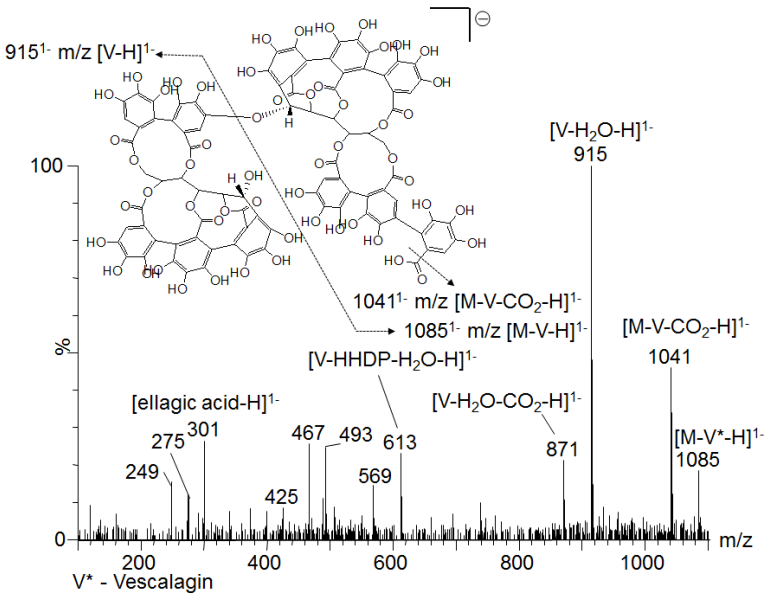


Figure 3.6: MS^E spectrum of a tentatively identified galloylated vescalagin dimer (**12**) obtained by HILIC with the proposed fragmentation scheme. V = vescalagin

The final oligomeric ellagitannin displaying multiple arrival times (**Figure 3.5A**) was tentatively identified as a glycosylated castalagin/vescalagin trimer (**18**), with the pentose moiety likely being either lyxose or xylose (similar to roburin B and C). (This compound is not visible in **Figure 3.5B** due to a lower sensitivity in HILIC). The mass spectrum of **18** shows both the doubly and triply charged ions at 1448 m/z $[M-2H]^{2-}$ and 965 m/z $[M-3H]^{3-}$, the former with two unique arrival times at 6.35 and 7.11 ms, respectively. **Figure 3.7** shows the high collision energy spectrum of the doubly charged species. The major fragment is a doubly charged ion at m/z 1382 representing the loss of a pentose ($[M\text{-pentose-}2H]^{2-}$). The product ions at 1065 m/z $[grandinin\text{-}H]^{1-}$ and 915 m/z $[vescalagin\text{-}H_2O\text{-}H]^{1-}$ confirm the occurrence of grandinin (vescalagin-glycoside) and vescalagin moieties in the molecule. Also the presence of a weak 631 m/z ion suggests that **18** may contain a castalagin unit.

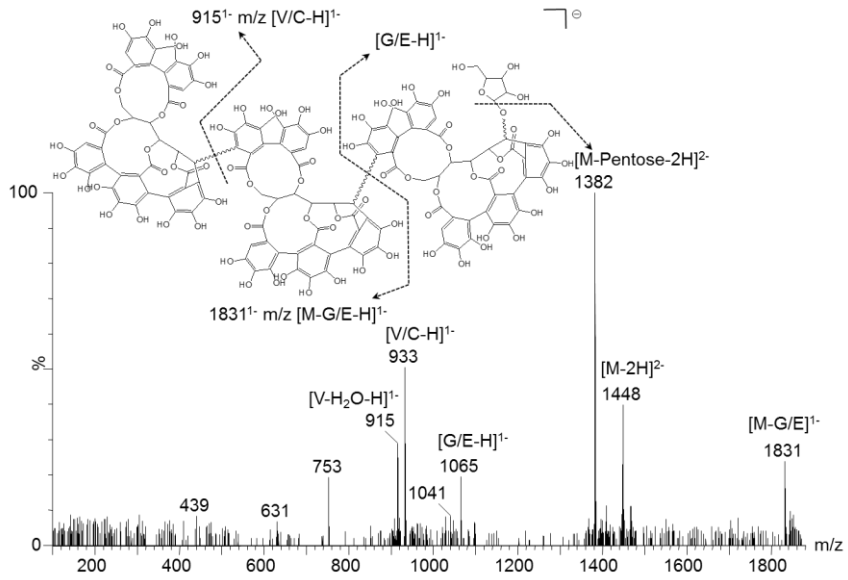


Figure 3.7: MS^E spectrum of a tentatively identified trimeric castalagin/vescalagin pentoside (**18**) with the proposed fragmentation scheme for this molecule. V/C = vescalagin/castalagin, G/E = grandinin/roburin E.

Comparing the HILIC and RP-LC for the analysis of oligomeric vescalagin/castalagin derivatives, co-elution of roburin A (**10ª**) and B (**11ª**) as well as the galloylated vescalagin dimer (**12**) with roburin C was observed in RP-LC, whereas in HILIC all

these species were resolved (**Figure 3.5**), making this the preferred chromatographic mode for these compounds.

The only remaining known vescalagin derivatives in chestnut are the pentoside derivatives grandinin (vescalagin-O-lyxose, **14^{a}**) and roburin E (vescalagin-O-xylose, **14^{b}**) [31]. The assignment of these compounds in RP-LC was based on the elution order reported by Herve du Penhoat *et al.* [30]. Two additional isomers with the same mass spectral properties (m/z 1065 $[M-H]^{1-}$, $C_{46}H_{33}O_{30}$) were detected at later RP-LC retention times; in contrast, only two peaks were detected in HILIC, likely due to co-elution. The fragmentation behaviour of the singly and doubly charged species of grandinin and roburin E were similar to those of vescalagin and castalagin, except for an additional fragment ion at 975/487 m/z ($[M-H]^{1-}$/$[M-2H]^{2-}$) likely resulting from intra-molecular fission of the pentose moiety (**Figure S9** and **S10**). A castalagin-vescalin/castalin dimer (**7**) and two di-hydrated derivatives of vescalagin (**5^{a}**) and castalagin (**5^{b}**) are tentatively identified in chestnut for the first time. The MS^{E} spectrum of the doubly charged molecular ion of **7** at 773 m/z $[M-2H]^{2-}$ showed fragment ions at 933 and 631 m/z, tentatively assigned as belonging to castalagin (**Figure S11**). Attempts to assign the castalin/vescalin unit by comparison of fragment ions with those vescalin (**1^{a}**) and castalin (**1^{b}**) (refer to **Table 3.1**) proved unsuccessful. The proposed structure and fragmentation scheme (**Figure S12**) are based on a C-C linked dimeric structure, as this is by far the most common structure observed for oligomeric ellagitannins.

The di-hydrated derivatives (**5^{a}** and **5^{b}**) were characterised by lower RP-LC and higher HILIC retention than vescalagin and castalagin, respectively, and molecular formulae for the singly charged molecular ions of $C_{41}H_{29}O_{28}$ (m/z 969.0880). The fragmentation spectra of the doubly charged molecular ion (**Figure 3.8B**) show the known ellagic acid fragments at m/z 301 and 275, confirming the presence of the HHDP group. An additional fragment in the MS^{E} spectrum of the singly charged molecular ion (**Figure 3.8A**) was detected at 667 m/z ($C_{27}H_{23}O_{20}$), which corresponds to a di-hydrated NHTP unit. Finally, despite partial co-elution with pedunculagin (**6**), a relatively clean UV spectrum obtained during HILIC analysis reveals a characteristic ellagitannin UV spectrum (**Figure 3.8C**).

Several galloylated ellagitannin derivatives, previously detected in oak [58, 59], are reported here in chestnut for the first time. Castavaloninic acid (**16^{b}**) and vescavaloninic acid (**16^{a}**) were assigned based on their RP-LC elution order [58] as well as HR-MS spectral information. In the case of vescavaloninic acid, ions at 1083 m/z [M-H_2O-H]$^{1-}$, 1065 m/z [M-2H_2O-H]$^{1-}$ and 1039 m/z [M-H_2O-CO_2-H]$^{1-}$ were observed, whereas for castavaloninic acid, the main fragment ion was observed at 1057 m/z [M-CO_2-H]$^{1-}$ (**Figure S13**) [56, 58].

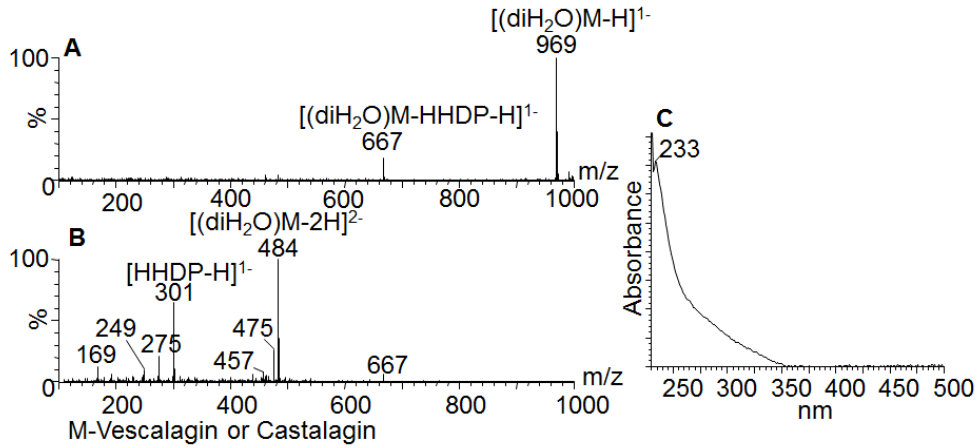


Figure 3.8: MS^{E} spectra of the (**A**) singly charged, and (**B**) doubly charged molecular ions of the proposed di-hydrated castalagin derivative (**5^{b}**). (**C**) shows the UV spectrum for this compound.

Compound **19** was tentatively identified as a galloylated vescalagin/castalagin derivative. The MS^{E} spectra of the doubly (542 m/z, [M-2H]$^{2-}$) and singly charged (1085 m/z [M-H]$^{1-}$) molecular ions are presented in **Figure S14**. The major fragment ions correspond to losses of CO_2 (1041 ([M-H]$^{1-}$) and 520 m/z ([M-2H]$^{2-}$)) [55], similar to castavaloninic and vescavaloninic acids, and of the intact galloyl group (915 and 457 m/z, [M-galloyl-xH]$^{x-}$). These data point to the galloyl group being attached via a phenolic OH. Furthermore, the presence of identical ellagic acid fragments as observed for castalagin/vescalagin implies that the galloyl group is positioned on the NHTP moiety. The related compound 1-O-galloyl castalagin has been identified in *Eugenia grandis* [22] by NMR, and was subsequently tentatively identified in chestnut

bark [55] based on nominal mass MS data. We propose that **19** is a similar derivative, but with the galloyl group attached via a phenolic OH as opposed to the carboxylic acid OH. Finally, based on the extent of water loss observed for this ion (large 457 m/z peak, *cf.* paragraph 3.2.1), it is hypothesised that this could be a vescalagin derivative. A second related galloylated derivative, compound **26** showed mass spectral properties consistent with a trihydroxybenzene substituted vescalagin/castalagin structure (**Figure S1**). Fragment ions for both the singly and doubly charged ions of **19** and **26** were identical, except for the additional ion at 1085 m/z corresponding to the molecular ion of **19**. A minor fragment ion at 739 m/z ($[M\text{-}HHDP\text{-}CO_2\text{-}H]^{1-}$ for **19**) in the MS^E spectra of both compounds points to the galloyl/trihydroxybenzene groups being coupled to the NHTP unit at the C1 position, as for 1-O-galloyl castalagin [22]. This hypothesis is supported by the absence of fragment ions at 409 and 453 m/z, which would indicate that the galloyl/trihydroxybenzene group is attached to the HHDP unit.

The six remaining ellagitannins detected in chestnut (compounds **13, 20, 24, 28, 29** and **32**), comprising in total 37 isomeric species, were all detected as doubly charged ions. These compounds all contain a vescalagin or castalagin unit coupled to an ellagitannin containing a cyclic glucose core (refer to **Table 3.2** below for further details on the latter compounds). Note that it is not possible to differentiate between castalagin and vescalagin derivatives in this class by MS, since the units are coupled via a C-C bond between the HHDP of the open core ellagitannin and the C-1 anomeric carbon of the closed core ellagitannin. Accordingly, the fragment ion at 915 m/z observed for all of these compounds results from cleavage of the C-C bond between the two constituent moieties. The mass fragment at 871 m/z ($[vescalagin\text{-}H_2O\text{-}CO_2\text{-}H]^{1-}$, *cf.* **Figure 3.4**) might, however, indicate the presence of a vescalagin unit in compounds **20, 28** and **29** (**Figure S15**).

These compounds are, therefore, distinguished based on the nature of the cyclic glucose ellagitannin unit, which may be one of pedunculagin (bis-HHDP glucopyranose, m/z 783, **13**), digalloyl glucose (m/z 483, **20**), tellimagrandin I (2,3-digalloyl-4,6-HHDP glucopyranose, m/z 785, **24**), tetragalloyl glucose (m/z 787, **28**), trigalloyl glucose (m/z 635, **29**) or ellagic acid (m/z 301, **32**). Since the singly charged closed core glucose moieties are present in the high collision energy spectra of all

compounds, it is relatively straightforward to assign these compounds based on IM-cleaned MS^E spectra (**Figure S16**). A large number of isomers for each of these species were resolved in RP-LC, whereas most isomers of the same derivative co-elute in HILIC. For example, sixteen isomers of the tentatively identified vescalagin/castalagin-digalloylglucose dimer (**20**) were resolved by RP-LC, in contrast to only one peak observed in HILIC.

Although it was not possible to obtain clear UV spectra for each of these derivatives due to co-elution, the available UV data (**Figure S17**) support the assignments discussed above. UV spectra for the castalagin/vescalagin-digalloylglucose (**20**) and -trigalloylglucose dimers (**29**) show the expected increase in the absorbance at 272-274 nm with an increase in the number of galloyl units [19]. The vescalagin/castalagin-pedunculagin dimer (**13**) shows a characteristic ellagitannin UV spectrum, i.e. reflecting presence of two HHDP units with no galloyl groups, whereas for castalagin/vescalagin-ellagic acid (**32**), a clear absorbance band at 374 nm confirms the presence of the ellagic acid group. The position of ellagic acid at the C1 position could be rationalised as either roburin A or D undergoing hydrolysis followed by lactonization.

Hydrolysis of vescalagin and castalagin results in the formation of vescalin (**1ª**) and castalin (**1ᵇ**), respectively, as well as ellagic acid (**36**) [2, 12]. The former products are characterised by UV spectra similar to vescalagin and castalagin, and similar simple fragmentation patterns (results not shown). Castalin and vescalin can be further dehydrated in acidic medium to produce vescalene (**21**) [5] (**Figure S1**). Three vescalene isomers were detected (**Table 3.1**), which show UV spectra similar to that of ellagic acid.

Finally, although strictly not derivatives of castalagin/vescalagin, several ellagic acid derivatives likely resulting from the hydrolysis of these compounds were detected, including ellagic acid (**36**), trimethoxy-ellagic acid (**38**), ellagic acid pentoside (**34**) and valoneic acid dilactone (**22**). Trimethoxy-ellagic acid (**38**) was identified based on accurate mass information (m/z 343.0461 $[M-H]^-$, $C_{17}H_{11}O_8$, 2 ppm) and fragment ions indicating the loss of three methyl groups, in accordance with the findings of Sanz *et al.* [32] for methoxylated acids and aldehydes. Ellagic acid pentoside (**34**, m/z 433 $[M-H]^{1-}$, $C_{19}H_{13}O_{12}$) showed a simple fragmentation spectrum indicating the cleavage of

the *O*-glycosidic bond. Similarly, the high collision energy mass spectra of valoneic acid dilactone (**22**), two isomers of which were detected, were characterised by relatively simple losses of CO_2 (m/z 425) and gallic acid (m/z 301 and 169).

3.3.2.2 Non-covalent cluster ions and adducts

Ellagitannins, like other oligomeric phenolics, are susceptible to the formation of non-covalent aggregates under ESI conditions [17, 45]. These adducts may complicate identification by ESI-MS. IM-MS can be beneficial in this regard by allowing identification of adducts and providing 'clean' isotopic information on the species of interest [63]. In the current study, several such aggregate species were detected in both RP-LC and HILIC. As an illustrative example, **Table S1** lists cluster ions observed for vescalagin and castalagin in both modes; the discussion below will focus on vescalagin (V), although identical behaviour was observed for castalagin. **Figure 3.9** shows the corresponding sections of the RP-LC×IM and HILIC×IM contour plots.

For both RP-LC and HILIC, dominant multiply charged dimeric $[2V\text{-}2H]^{2-}$ and trimeric $[3V\text{-}2H]^{2-}$ clusters were detected; a pentameric $[5V\text{-}3H]^{3-}$ cluster was detected in RP-LC only. Inspection of the low energy spectra obtained from the chromatographic trace reveals a mass offset (ΔDa) of 0.5 for the isotope peaks due to the presence of doubly charged dimeric cluster at 933 m/z ($[2V\text{-}2H]^{2-}$, **Figure S18A**). However, singly and doubly charged ions are resolved by IM, with arrival times of 6.62 and 3.11 ms, respectively.

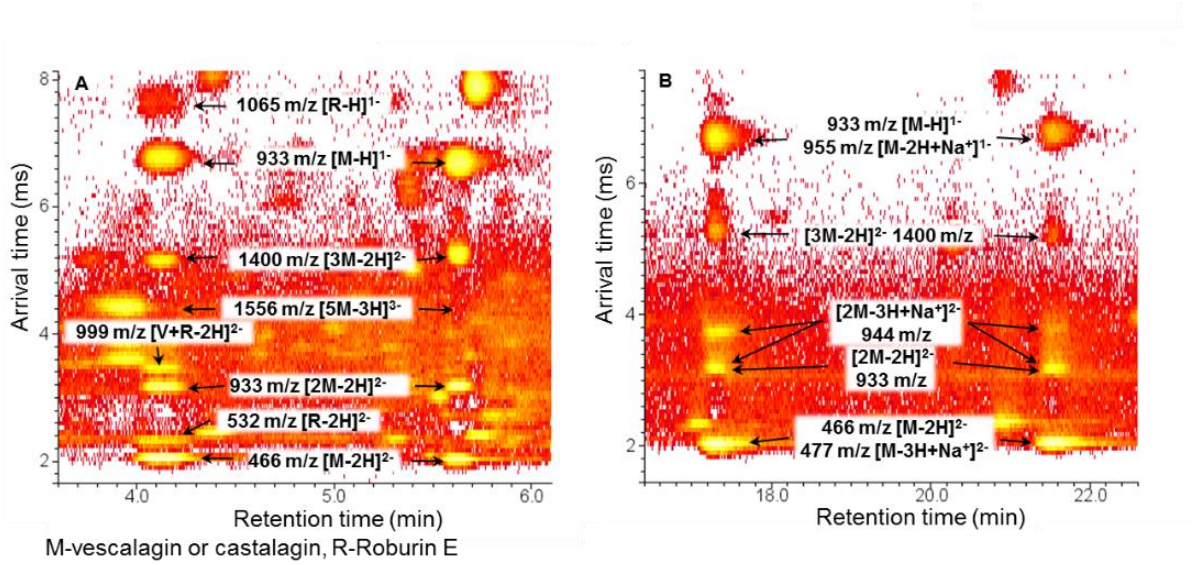


Figure 3.9: Illustration of the non-covalent adducts of vescalagin (**17[a]**) and castalagin (**17[b]**) obtained for (**A**) RP-LC- and (**B**) HILIC-ESI-IM-HR-MS. V = vescalagin, C = castalagin, R = roburin E.

This allows the option of filtering MS data according to arrival time to obtain mass spectra for selected target species only (illustrated in **Figure S18B** and **C**). The mass spectra of the trimeric doubly charged ($[3V\text{-}2H]^{2-}$, 1400 m/z, t_d 5.11 ms) and pentameric triply charged ($[5V\text{-}3H]^{3-}$, 1556 m/z, t_d's 4.07 and 4.28 ms) vescalagin cluster ions are shown **Figure S19**. Interestingly, two pentameric cluster ions partially separated by IM were detected for both vescalagin and castalagin by RP-LC. Similarly, the sodiated doubly charged dimeric vescalagin and castalagin ionic species ($[2V/C\text{-}3H+Na^+]^{2-}$, 944 m/z) detected in HILIC also showed different arrival times at 3.17 and 3.66 ms (**Figure 3.9, Table S1**).

A closer look at the mass spectrum of the doubly charged molecular ion of vescalagin (m/z 466, t_d=1.93 ms, **Figure 3.10**) reveals the presence of multiply charged clusters containing two or four vescalagin units. Several isotope ion series are also present in **Figure 3.10**, labelled using the nomenclature $[x,^{13}C(y)]^{z-}$, where x represents the number of vescalagin ions, y the number of ^{13}C atoms and z the charge of the cluster/ion (z = 2x). The main peaks in this spectrum represent the M+1 isotopologues of the doubly charged molecular ion, as confirmed by their mass offset (ΔDa) of 0.501

Da. Two additional m/z envelopes defined by mass offsets of 0.130 and 0.250 Da are due to dimeric and tetrameric clusters of vescalagin with 4 and 8 charges, respectively. For example, m/z 466.1539 ([4,^{13}C(1)]$^{8-}$) represents a octuply charged cluster containing four doubly charged vescalagin units with one ^{13}C atom (the mass offset of 0.1250 Da corresponds well with the predicted offset of 0.1254 Da). **Table S2** presents a summary of the heteroisotopic clusters detected for vescalagin with mass accuracy values. This type of cluster behaviour was mainly observed for vescalagin and castalagin due to their tendency for clusters formation, as well as their natural abundance in the sample.

Of particular interest is the occurrence of mixed cluster ions resulting from the co-elution of different species. As an example, vescalagin (**17[a]**) partially co-elutes with roburin E (**14[b]**) in RP-LC. In the mass spectrum obtained from the chromatographic trace of vescalagin, a doubly charged mixed vescalagin-roburin E dimeric cluster is detected at m/z 999 [Roburin E+V-2H]$^{2-}$ (**Figure 3.11**). This adduct has a unique arrival time of 3.38 ms, and is not observed for castalagin or vescalagin in HILIC (**Figure 3.9B**).

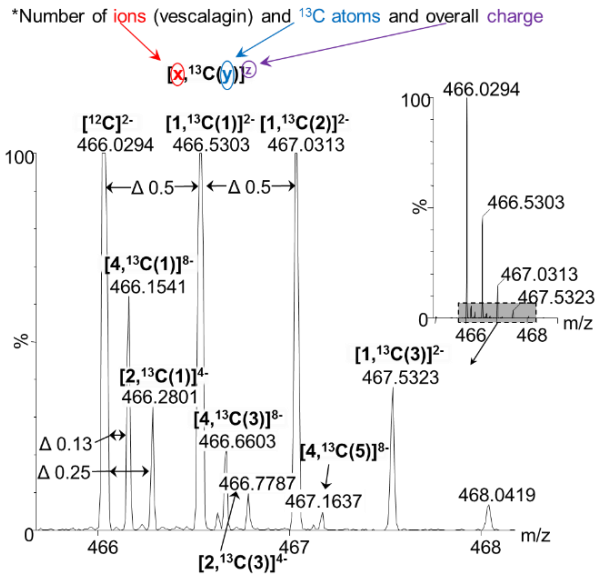


Figure 3.10: Close up of the mass spectrum of the doubly charged molecular ion of vescalagin illustrating the presence of multiply charged heteroisotopic cluster ions. *The legend defines nomenclature used in the assignment of m/z peaks in the spectrum.

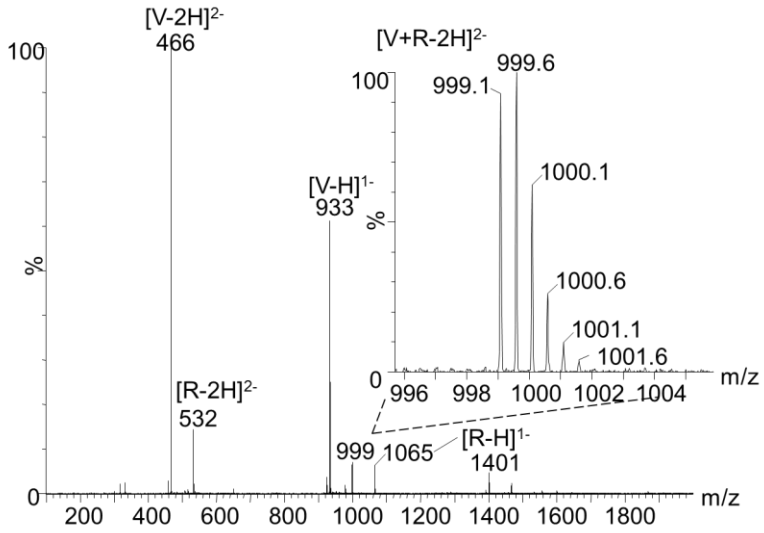


Figure 3.11: Mass spectrum obtained from the RP-LC chromatographic trace at 4.12 minutes for vescalagin (V, **17[a]**), illustrating the formation of a mixed doubly charged dimeric cluster comprising vescalagin and the co-eluting roburin E (R, **14[b]**) at 999 m/z ($[V+R-2H]^{2-}$).

The mass spectra of hydrolysable tannins in HILIC differed from those under RP-LC conditions in that sodium adducts with the general formula $[xM-yH+Na^+]^{1-y}$ were observed (with the exceptions of monogalloyl glucose (**2**), ellagic acid-pentoside (**3**), gallic acid (**9**) and trimethoxy-ellagic acid (**38**)). The reasons for these discrepancies are likely related to the mobile phases used in the RP-LC and HILIC, where in the former case the excess water could effectively solvate sodium ions. For example, vescalagin produces two doubly charged (477 m/z, $[V-3H+Na^+]^{2-}$ and 944 m/z, $[2V-3H+Na^+]^{2-}$) and one singly charged (955 m/z, $[V-2H+Na^+]^{1-}$) sodium adducts (**Figure S20**). The arrival times for the singly charged deprotonated (933 m/z, $[V-H]^{1-}$) and sodiated (m/z 955, $[V-2H+Na^+]^{1-}$) species were identical (6.62 ms), as were those of the corresponding doubly charged (466 m/z, $[V-2H]^{2-}$ and 477 m/z, $[V-3H+Na^+]^{2-}$) ions (1.93 ms). Interestingly, however, the sodiated doubly charged dimeric cluster at m/z 944 ($[2V-3H+Na^+]^{2-}$) did show multiple arrival times at 3.17 and 3.66 ms (illustrated in **Figure 3.9**).

3.3.2.3 Ellagitannins with a closed glucose core

Table 3.2 and **Figure 3.12** summarise the ellagitannins with a cyclic glucose core tentatively identified in chestnut. Of the eight species, which comprise in total 29 isomeric structures, only tellimagrandin II (trigalloyl-HHDP-glucose, **35**) and several isomers of tellimagrandin I (digalloyl-HHDP-glucose, **27**) have previously been identified in chestnut [32, 55]. All the compounds listed in **Table 3.2** are derivatives of HHDP-glucose (**3**), the simplest of the closed glucose ellagitannins in chestnut. It is also worth noting that the arrival times for all isomers listed in **Table 3.2** were identical (for the same ion species); IM therefore provided no additional separation for this class of compounds. RP-LC was found to provide better separation for the closed glucose ellagitannins, with a total of 27 isomers resolved compared to 20 in HILIC.

The MS^E spectra of two HHDP-glucose (**3**) isomers as well as the proposed fragmentation pathway are presented in **Figure S21**. The primary fragment ions correspond to the loss of a glucose moiety, leading to formation of an ellagic acid ion at m/z 301, followed by dehydration (m/z 275) (*cf.* **Figure S2**). For one of the isomers, an additional fragment ion at 421 m/z could be due to the intra-molecular fission of the sugar moiety. This fragment ion can be used to distinguish two groups of isomers: isomers **3^c** and **3^d** featuring this fragment ion co-eluted at 1.17 min in RP-LC, but were resolved (11.80 and 13.32 min) in HILIC.

Compounds **6^{a-e}** represent pedunculagin (di-HHDP-glucose) isomers (5 of these were separated in both RP-LC and HILIC). MS^E spectra of these compounds show, in addition to the ellagic acid ions, a 481 m/z fragment ion resulting from the loss of a single HHDP unit to form $[HHDP\text{-}glucose\text{-}H]^{1-}$. Compound **4** was tentatively identified as a decarboxylated derivative of pedunculagin ($C_{34}H_{24}O_{20}$, 3.25 ppm). However, neither MS^E nor UV data were sufficient to allow further deductions on the structure of this compound to be made.

For galloyl-HHDP-glucose (**8**, $C_{27}H_{22}O_{18}$, m/z 633 $[M\text{-}H]^{1-}$), only four isomers were resolved in HILIC, whereas 11 were separated by RP-LC. The dominant fragment ions for these species correspond to losses of galloyl (m/z 481), HHDP (m/z 331) and galloyl-glucose moieties (m/z 301) as well as the formation of gallic acid (m/z 169). The galloylation site could be determined for some isomers based on the MS^E spectra

obtained in HILIC: for the isomer eluting at 9.11 min in HILIC, the fragment ion at 421 m/z confirms the galloyl group to be substituted at the 6 position of the glucose core (**Figure S22**).

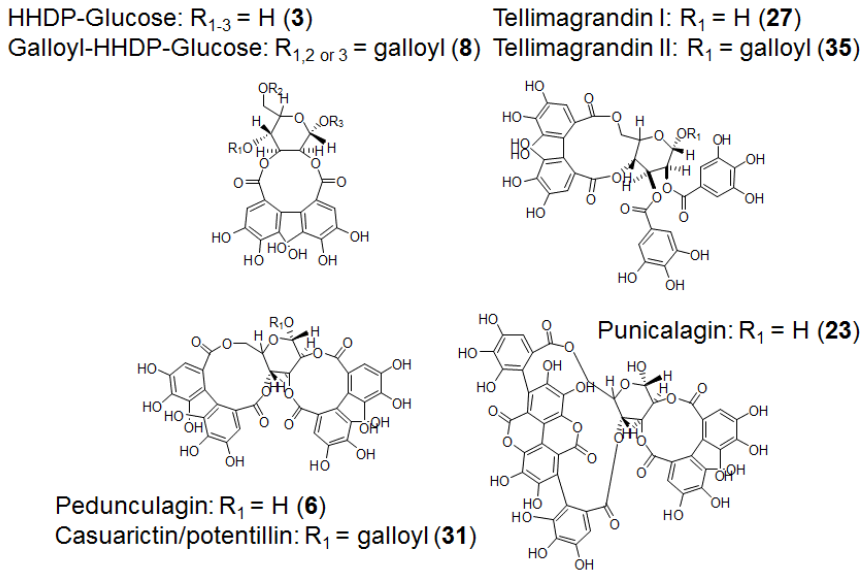


Figure 3.12: Structures of ellagitannins containing a cyclic glucose core tentatively identified in chestnut. Numbers in brackets correspond to **Table 3.2.**

Compound **23**, predominately detected as its doubly charge ion at 541 m/z $[M-2H]^{2-}$, was tentatively identified as punicalagin (2,3-HHDP-4,6-gallagyl-glucose, $C_{48}H_{28}O_{30}$ (1.66 ppm) based on HR-MS data. Only one isomer of this compound was detected in both HILIC and RP-LC.

For tellimagrandin I (digalloyl-HHDP-glucose, **27**), five and three isomers were separated by RP-LC and HILIC, respectively. MS^E spectra obtained in HILIC proved more informative for these compounds, since more extensive fragmentation was observed: fragment ions corresponding to $[M\text{-galloyl-H}]^{1-}$ (m/z 633), $[M\text{-HHDP-H}]^{1-}$ (m/z 483), $[M\text{-HHDP-galloyl-H}]^{1-}$ (m/z 331), $[\text{ellagic acid-H}]^{1-}$ (m/z 301) and gallic acid (m/z 169) were detected (**Figure S23**). No notable differences in the fragmentation of isomers to allow their differentiation were observed.

Compound **31** was tentatively identified as casuarictin/potentillin (1-O-galloyl-bis-HHDP-glucose, these compounds only differ in terms of α/β configurations of the glucose core, respectively). The MS^E spectrum for this species, showed fragment ions of 633 m/z $[M\text{-}HHDP\text{-}H]^{1-}$ and 481 m/z $[M\text{-}HHDP\text{-}galloyl\text{-}H]^{1-}$. Finally, the MS^E spectrum of compound **35**, **Figure S24**, showed fragment ions of 767 $[M\text{-}gallate\text{-}H]^{1-}$,467 $[M\text{-}HHDP\text{-}gallate\text{-}H]^{1-}$, 301 $[ellagic\ acid\text{-}H]^{1-}$ and a dominant 169 $[gallic\ acid\text{-}H]^{1-}$ ion, tentatively identifying this compound as tellimagrandin II (trigalloyl-HHDP-glucose).

Although not discussed here, online UV spectra support the above assignments based on MS data. In particular, the relative intensity of the absorbance band at 273-278 nm in relation to the ratio of galloyl to HHDP units [19] proved useful in this regard (illustrated in **Figure S25** for HHDP-glucose (**3**) and galloyl-HHDP-glucose (**8**) as example.

Table 3.2: Closed glucose core ellagitannins detected in chestnut using RP-LC- and HILIC-IM-HR-MS.

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
HHDP-Glucose (3[a-d])	$C_{20}H_{18}O_{14}$	(1.00, 1.41)[a+b]	9.27[a]	230.00	481.0636	3.67	3.04	197	nd*				300.9989, 275.0200
			9.64[b]										300.9989, 275.0200
		1.17[cd]	11.80 [c]										421.0421, 300.9993, 275.0198
			13.32[d]										421.0421, 300.9993, 275.0198
Pedunculagin - CO_2 (4)	$C_{33}H_{24}O_{20}$	1.12	20.26	co-elution[a]	739.0807	3.25	4.97	251	nd				695.0921,677.0807, 631.0611, 613.0488, 481.0638, 425.0145, 399.0362, 355.0304, 191.0191
Pedunculagin I (6[a-e])	$C_{34}H_{24}O_{22}$	1.26	25.80	co-elution	783.0687	0.76	5.18	257	391.0322	3.84	1.66	284	739.0815, 695.0923, 677.0807, 613.0481, 481.0639,425.0148, 399.0360
		1.91	16.40										299.0179
		3.75	11.77										481.0625, 300.9973, 275.0191
		5.18	10.77										co-elution*
		6.95	10.15										633.0750, 613.0483, 483.0784, 300.9989, 275.0178
Galloyl-HHDP-glucose (8[a-k])	$C_{27}H_{22}O_{18}$	1.67	8.16, 8.40, 9.11, 10.06[#]	273	633.0728	0.01	3.86	222	nd			nd	co-elution
		2.59											co-elution
		4.67											300.9962, 275.0178
		5.84											co-elution
		5.93											co-elution
		6.22											481.0820, 300.9983
		6.61											co-elution
		6.95											co-elution
		7.13											co-elution
		8.38											301.0004, 275.0204
		8.72											co-elution
Punicalagin (23[a])	$C_{48}H_{28}O_{30}$	6.74	23.46	230[sh]	1083.0605	1.66	7.80	315	541.0269	2.59	2.35	338	co-elution
Tellimagrandin I (27[a-d])	$C_{34}H_{26}O_{22}$	7.84	8.73, 8.97, 9.73[#]	276	785.0842	0.64	5.45	263	392.0338	0.00	1.66	286	483.0783, 300.9990, 275.0190
		9.26											483.0777, 300.9980, 275.0193
		9.78											nd
		10.27											300.9985
		10.81											nd

Table 3.2 (continued).

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
Casuarictin or potentillin (31)	$C_{41}H_{28}O_{26}$	8.68	12.95	nd	935.0798	0.75	6.62	290	467.0377	-7.71	1.93	306	nd
Tellimagrandin II (35)	$C_{41}H_{30}O_{26}$	11.23	18.20	nd	937.0977	3.20	6.69	292	468.0455	4.27	2.00	312	787.1030, 300.9986, 275.0165, 169.0134, 137.0250, 125.0205

* t_d: arrival time; nd: not detected; co-elution: no clear data obtained due to co-elution.

Assignment of corresponding isomers between RP-LC and HILIC was not possible due to fewer isomers detected in HILIC

3.3.2.4 Gallotannins

Table 3.3 summarises the gallotannins identified in the chestnut extract, which include mono-(**2**), di-(**15**), tri-(**30**), tetra-(**33**) and penta-galloylglucose (**37**). UV spectra for these compounds are identical, with absorbance maxima at 272-277 nm representing a bathochromic shift compared to gallic acid [57]. Representative high energy mass spectra for each of these classes are presented in **Figure S26**. Fragmentation of these molecules involves cleavages resulting in charged gallic acid (m/z 169) and trihydroxybenzene (m/z 125) moieties, neutral loss(es) of galloyl group(s) ([M-152]$^{-}$), as well as intramolecular $^{0,2}A^{-}$ and $^{2,4}A^{-}$ cleavages of glucose moieties according to the nomenclature of Domon and Costello [64] and additional losses of water for the higher MW compounds (**Figure S27**). Sodium adducts [M-2H+Na^{+}]$^{1-}$ were also detected in HILIC, for example at m/z 657 and 809 for tri- and tetra-galloyl glucose species, respectively.

In addition to the gallotannins, free gallic acid (**9**) and digallic acid (**25**) were also detected. The UV spectrum for digallic acid (**25**) showed a characteristic shoulder around 300 nm indicating the presence of a depsidic bond [57].

Regarding the number of isomers resolved by RP-LC and HILIC (**Table 3.3**), the selectivity (in terms of number of isomers resolved) of RP-LC increases relative to HILIC with the degree of galloylation. For example, for monogalloyl glucose, thirteen isomers were resolved by HILIC compared to only six by RP-LC, whereas eleven isomers of trigalloyl glucose were separated by RP-LC compared to four in HILIC. For a complete analysis of galloyl glucose species in chestnut, comprehensive HILIC×RP-LC separation would therefore be required.

Table 3.3: Gallotannin species detected in chestnut by RP-LC and HILIC coupled to ion-mobility and high resolution mass spectrometry.

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{1-}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions
Monogalloyl glucose (2)	$C_{13}H_{16}O_{10}$	0.92, 1.46, 1.87, 2.07, 2.35, 3.13, 3.33, 4.32	6.50, 8.02, 8.59, 8.80, 9.09, 9.41, 9.64, 10.14, 11.29, 13.78, 14.70, 15.39, 17.11	273	331.0664	-0.38	2.21	168	nd*				271.0512, 211.0283, 169.0169, 124.0177
Gallic acid (9)	$C_7H_6O_5$	2.03	2.95	270	169.0131	-3.55	0.83	105	nd				125.0227, 124.0143
Digalloyl glucose (15)	$C_{20}H_{20}O_{14}$	3.21, 5.53, 5.89, 6.13, 6.36, 6.93, 7.47, 7.85, 7.99, 8.23, 9.58	6.13, 6.35, 6.82, 7.08, 7.47, 7.57, 7.78, 7.93, 8.36, 9.23	279	483.0776	0.24	2.90	191	241.0353	1.66	0.97	228	331.0759, 313.0629, 271.0510, 211.0284, 193.0262, 169.0168, 125.0259
Digallic acid (25)	$C_{14}H_{10}O_9$	6.89, 8.13	3.38	276	321.0253	1.87	2.00	160	nd				169.0251, 125.0308
Trigalloyl glucose (30)	$C_{27}H_{24}O_{18}$	8.46, 8.88, 9.00, 9.12, 9.45, 9.81, 9.90, 10.10,10.17 10.49, 10.67, 10.95	6.76, 6.95, 7.42, 7.57	276	635.0886	0.24	3.86	220	317.0406	0.95	1.38	261	483.1064, 465.0943, 313.0766, 271.0618, 211.0365, 193.0262, 169.0243
Tetragalloyl glucose (33)	$C_{34}H_{28}O_{22}$	10.45, 11.26, 11.34, 11.43	7.78, 8.57	279	787.0991	-0.38	5.52	264	393.0450	-2.04	1.73	297	635.1309, 617.1185, 483.1060, 465.0948, 301.0845, 169.0223
Pentagalloyl glucose (37)	$C_{41}H_{32}O_{26}$	11.64	9.07	280	939.1111	0.75	6.76	292	469.0506	-1.49	2.14	325	769.1335, 633.1162, 483.1089, 481.0923, 331.0874

* t_d: arrival time; nd: not detected.

Plots of retention time vs. arrival time for HILIC and RP-LC, **Figure 3.13**, illustrate identical arrival times for galloyl glucose species of the same degree of galloylation. Also evident from this figure, is that RP-LC retention increased with the degree of galloylation, which effectively results in a size separation for this compound class. This is in contrast to HILIC, where extensive co-elution of compounds of different degrees of galloylation is observed. Another interesting feature evident from **Figure 3.13** is that similar arrival times are observed for the galloyl glucose compounds and the corresponding HHDP-glucose derivatives containing two less galloyl groups (for example **30** and **8**, **33** and **27**, **37** and **35**), indicating that the HHDP group contributes roughly the same to the $^{TW}CCS_{N2}$ values of hydrolysable tannins as do two galloyl groups.

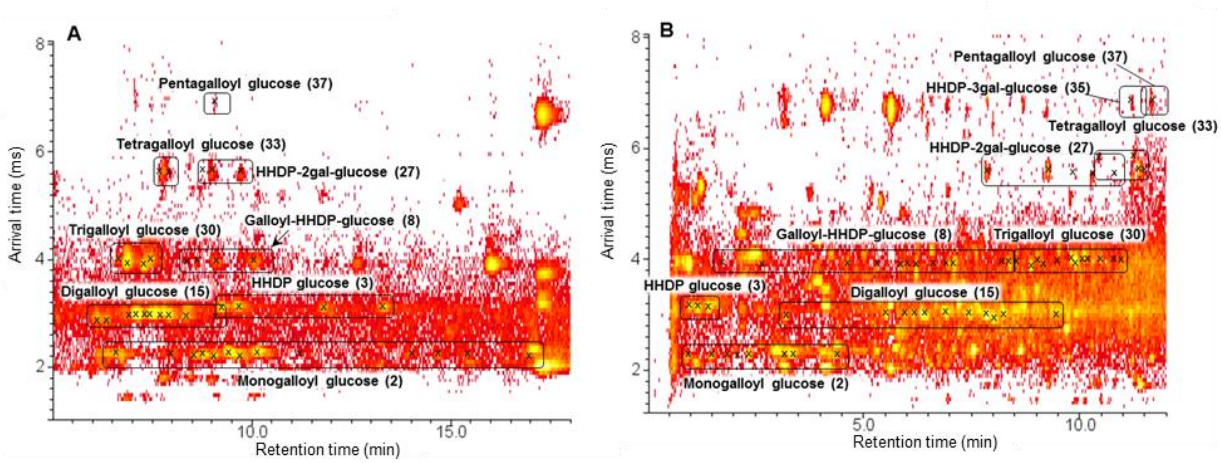


Figure 3.13: HILIC×ion mobility (**A**) and RP-LC×ion mobility (**B**) TIC (m/z 100-2000) contour plots obtained for the gallotannins in chestnut. Refer to **Tables 3.2** and **3.3** for further details on compound assignment.

3.3.3 Summary of hydrolysable tannins identified in chestnut

A summary of the molecular species and their respective numbers of isomers and conformers detected in the chestnut extract is presented in **Table 3.4**. The majority of these are ellagitannins, of which open glucose core vescalagin and castalagin derivatives (**Table 3.1**) were predominant. Of the thirty-one ellagitannin species

detected, nineteen are tentatively identified in chestnut for the first time in the current work. The large number of compounds identified in the current work compared to previous reports on chestnut tannins which utilised MALDI-TOF analysis [18, 39, 40, 65, 66] can be ascribed to the combination of (1) efficient chromatographic separation using two complementary modes, (2) the availability of high resolution mass spectral data which allowed molecular formula determination, and (3) high collision energy MS data, which together with the discussed fragmentation pathways strengthens tentative structural assignment. Clearly, NMR would be required to confirm the proposed structures of the new tentatively assigned compounds – including their stereochemistry. Nevertheless, the data reported here indicate several potentially interesting novel compounds in chestnut, and upscaled HILIC and RP-LC methods might prove useful for fractionation purposes.

The total number of isomers detected for gallotannins and some closed core ellagitannins exceed the number of isomers resolved by either RP-LC or HILIC, which points to important selectivity differences between these two separation modes on the one hand, and highlights the importance of complementary separation modes for these compounds on the other. In total, 128 and 90 isomeric species were detected using RP-LC and HILIC, respectively, with each separation mode providing better separation for particular classes of hydrolysable tannins.

Finally, six dimeric and trimeric vescalagin/castalagin conformers were resolved by ion mobility. Conformers were detected for roburins A-D, as well as for newly identified C-galloylated vescalagin (**12**) and glycosylated vescalagin/castalagin trimeric (**18**) compounds; the conformers of all these compounds showed identical retention times in both HILIC and RP-LC

Table 3.4: Summary of the molecular species and numbers of isomers detected in chestnut in the current work by RP-LC- and HILIC-IM-HR-MS.

		Isomers resolved			Conformers resolved
Compound class	Molecular species	HILIC	RP-LC	Number of isomers	Ion mobility
Vescalagin and castalagin derivatives	23 (14)*	39	67	67 (53)*	6
Cyclic glucose core ellagitannins	8 (5)*	20	27	28 (23)*	0
Gallotannins	7 (1)*	31	34	41 (2)*	0
Total	38 (20)*	90	128	136 (78)*	6

* numbers in brackets indicate molecular species/isomers tentatively identified for the first time in chestnut.

3.3.4 Measured $^{TW}CCS_{N2}$ values of chestnut hydrolysable tannins and comparison with drift tube IM data

$^{TW}CCS_{N2}$ values for the identified compounds (**Tables 3.1-3.3**) were derived using an established calibration protocol with poly-DL-alanine as calibrant (**Section 3.2.6**). Since to our knowledge these are the first reported $^{TW}CCS_{N2}$ values for hydrolysable tannins, we endeavoured to confirm the validity of our data by comparison of $^{DT}CCS_{N2}$ values for selected hydrolysable tannins with those measured on a drift tube IM spectrometry (DTIMS) instrument. Unlike TWIMS, $^{DT}CCS_{N2}$ values can be directly determined from measured arrival times in DTIMS without the need for calibration using reference compounds. The results are summarised in **Table S3**, where it is clear that generally good agreement was obtained between the two instruments, supporting the validity of CCS values reported for hydrolysable tannins in **Tables 3.1-3.3.**

$^{TW}CCS_{N2}$ values measured for hydrolysable tannins varied between ~100 Å^2 for gallic acid and ~700 Å^2 for some adducts, in the general sequence hydrolysed ellagitannins < gallotannins < closed glucose core ellagitannins < castalagin/vesgalagin derivatives < adducts. In general, $^{TW}CCS_{N2}$ values for isomeric compounds were near identical. Notable exceptions are where multiple conformers were detected for oligomeric vescalagin/castalagins (**10-12, 18**), as well as the sodium adducts of doubly charged dimeric clusters ($[2M-3H+Na^+]^{2-}$) of both vescalagin and castalagin, which showed multiple arrival times.

3.4 Conclusions

In this study we have exploited the complementary information provided by RP-LC and HILIC separations coupled to ion mobility and low and high collision energy HR-MS for the in-depth analysis of chestnut hydrolysable tannins. A total of 38 molecular species, comprising in total 136 isomers, were detected, including 20 species (78 isomers) tentatively identified for the first time in chestnut. A further six oligomeric vescalagin and castalagin oligomeric conformers were resolved by ion mobility. Ion mobility also offered the advantage of allowing the obtainment of 'clean' mass spectra by filtering the data according to arrival time, which significantly improved data quality and interpretation.

The large number of isomeric hydrolysable tannin species places a premium on chromatographic separation to allow their tentative identification by MS. Particular classes of hydrolysable tannins were better separated by either HILIC or RP-LC. For example, several vescalagin/castalagin galloyl glucose oligomeric isomers were resolved in RP-LC but co-eluted in HILIC, whereas HILIC was superior for the separation of vescalagin/castalagin monomers and dimers (roburins A-D).

Clearly, further research is required for the unambiguous structural elucidation of compounds tentatively identified in the current work. Nevertheless, the proposed methodology has proved useful in highlighting the hitherto unknown complexity of chestnut tannin composition. Considering also the data reported in the second part of this work, which focuses on gallotannins in tara (chapter 4) the combination of RP-LC and HILIC separations with ion mobility and HR-MS shows promise as a powerful methodology for the analysis of hydrolysable tannins in a range of natural products.

3.5 References

1. Quideau S, Deffieux D, Douat-Casassus C, Pouysegu L (2011) Natural Products Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. Angew Chem 50:586–621. doi: 10.1002/anie.201000044

2. Jurd L (1956) Plant Polyphenols. I. The Polyphenolic Constituents of the Pellicle of the Walnut. J Am Chem Soc 78:3345–3448. doi: 10.1021/ja01595a050

3. White T (1957) Tannins-Their occurrence and significance. J Sci Food Agric 8:377–385. doi: 10.1002/jsfa.2740080702

4. Okuda T, Yoshida T, Hatano T (1990) Oligomeric hydrolysable tannins, a new class of plant polyphenols. Heterocycles 30:1195–1218

5. Quideau S, Jourdes M, Lefeuvre D, Montaudon D, Saucier C, Glories Y, Pardon P, Pourquier P (2005) The chemistry of wine polyphenolic C-glycosidic ellagitannins targeting human topoisomerase II. Chem - A Eur J 11:6503–6513. doi: 10.1002/chem.200500428

6. Sánchez-González C, Ciudad CJ, Noé V, Izquierdo-Pulido M (2014) Walnut polyphenol metabolites, urolithins A and B, inhibit the expression of the prostate-specific antigen and the androgen receptor in prostate cancer cells. Food Funct 5:2922–30. doi: 10.1039/c4fo00542b

7. Asanaka M, Kurimura T, Koshiura R, Okuda T, Mori M, Yokoi H (1988) Tannins as candidates for anti-HIV drug. Int J Immunopharmacol 10:35

8. Kashiwada Y, Huang L, Kilkuskie RE, Bodner AJ, Lee K-H (1992) New hexahydroxydiphenoyl derivatives as potent inhibitors of HIV replication in H9 lymphocytes. Bioorg Med Chem Lett 2:235–238. doi: 10.1016/S0960-894X(01)81071-4

9. Puppala M, Ponder J, Suryanarayana P, Reddy GB, Petrash M, LaBarbera D V. (2012) The isolation and characterization of b-glucogallin as a novel aldose reductase inhibitor from emblica officinalis. PLoS One 7:1–9. doi: 10.1371/journal.pone.0031399

10. Schmidt OJ, Mayer W (1956) Natürliche Gerbstoffe. Angew Chem 68:103–115

11. Herve du Penhoat CLM, Michon VMF, Ohassan A, Peng S, Scalbert A, Gage D (1991) Roburin A, A dimeric ellagitannin from heartwood of Quercus robur. Phytochemistry 30:329–332. doi: 10.1016/0031-9422(91)84148-L

12. Schmidt OT, Blinn F, Lademann R (1952) Über die Bindung der Ellagsäure in Corilagin und Chebulagsäure. XII. Mitteilung über natürliche Gerbstoffe. Justus Liebigs Ann Chem 576:75–84

13. Okuda T, Yoshida T, Hatano T (1989) New Methods of Analyzing Tannins. J Nat Prod 52:1–31. doi: 10.1021/np50061a001

14. Haslam E (1986) Secondary Metabolism - Fact and Fiction. Nat Prod Rep 3:217–249

15. Haslam E, Cai Y (1994) Plant polyphenols (vegetable tannins): Gallic acid metabolism. Nat Prod Rep 11:41–66. doi: 10.1039/np9941100041

16. Quideau S, Feldman KS (1996) Ellagitannin chemistry. Chem Rev 96:475–504. doi: 10.1021/jo034495x

17. Moilanen J, Sinkkonen J, Salminen JP (2013) Characterization of bioactive plant ellagitannins by chromatographic, spectroscopic and mass spectrometric methods. Chemoecology 23:165–179. doi: 10.1007/s00049-013-0132-3

18. Pasch H, Pizzi A (2002) Considerations on the macromolecular structure of chestnut ellagitannins by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. J Appl Polym Sci 85:429–437. doi: 10.1002/app.10618

19. Salminen JP, Ossipov V, Loponen J, Haukioja E, Pihlaja K (1999) Characterisation of hydrolysable tannins from leaves of Betula pubescens by high-performance liquid chromatography-mass spectrometry. J Chromatogr A 864:283–291. doi: 10.1016/S0021-9673(99)01036-5

20. Mueller-Harvey A (2001) Analysis of hydrolysable tannins. Anim Feed Sci Technol 91:3–20

21. Mayer W, Gabler W, Riester A, Korger H (1967) Über die Gerbstoffe aus dem

Holz der Edelkastanie und der Eiche, II. Die Isolierung von Castalagin, Vescalagin, Castalin und Vescalin. Liebigs Ann Chem 707:177–181

22. Nonaka G, Ishimaru K, Watanabe M, Nishioka I, Yamauchi T, Wanc ASC (1987) Tannins and Related Compounds. LI. Elucidation of the stereochemistry of the triphenoyl moiety in castalagin and vescalagin, and isolation of 1-O-galloyl castalagin from Eugenia grandis. Chem Pharm Bull.hem Pharm 35:217–220

23. Nonaka G, Sakai T, Tanaka T, Mihashi K, Nishioka I (1990) Tannins and related compounds. XCVII. Structure revision of C-glycosidic ellagitannins, castalagin, vescalagin, casuarinin and stachyurin, and related hydrolyzable tannins. Chem Pharm Bull 38:2151–2156. doi: 10.1248/cpb.38.2151

24. Matsuo Y, Wakamatsu H, Omar M, Tanaka T (2015) Reinvestigation of the Stereochemistry of the C- Glycosidic Ellagitannins, Vescalagin and Castalagin. Org Lett 17:46–49

25. Okuda T, Mori K, Seno K, Hatano T (1979) Constituents of Geranium thunbergII Sieb. et Zucc. J Chromatogr 171:313–320. doi: 10.1016/S0021-9673(01)95310-5

26. Okuda T, Yoshida T, Hatano T, Yazaki K, Ashida M (1980) Ellagitannins of the casuarinaceae, stachyuraceae and myrtaceae. Phytochemistry 21:2871–2874. doi: 10.1016/0031-9422(80)85058-8

27. Okuda T, Yoshida T, Hatano T, Yazaki K, Kira R, Ikeda Y (1986) Chromatography of Tannins II. Preparative Fractionation of Hydrolyzable Tannins By Centrifugal Partition Chromatography. J Chromatogr chromatograhy 362:375–381

28. Hatano T, Yoshida T, Okuda T (1988) Multiple peaks in high performance liquid chromatography of some hydrolyzable tannins. J Chromatogr 435:285–295

29. Scalbert A, Duval L, Peng S, Monties B, Du Penhoat C (1990) Polyphenols of Quercus robur L. IIa. Preparative isolation by low-pressure and high-pressure liquid chromatography of heartwood ellagitannins. J Chromatogr 502:107–119

30. Herve du Penhoat CLM, Michon VMF, Peng S, Viriot C, Scalbert A, Gage D

(1991) Structure elucidation of new dimeric ellagitannins from Quercus robur L. roburins A-E. J Chem Soc, Perkin Trans 1653–1660

31. Viriot C, Scalbert A, Hervé du Penhoat CLM, Moutounet M (1994) Ellagitannins in woods of sessile oak and sweet chestnut dimerization and hydrolysis during wood ageing. Phytochemistry 36:1253–1260. doi: 10.1016/S0031-9422(00)89647-8

32. Sanz M, Cadahia E, Esteruelas E, Munoz AM, Fernandez De Simon B, Hernandez T, Estrella I, Cadahia E, Esteruelas E, Munoz AM, Fernandez De Simon B, Hernandez T, Estrella I (2010) Phenolic compounds in chestnut (Castanea sativa Mill.) heartwood. Effect of toasting at cooperage. J Agric Food Chem 58:9631–9640. doi: 10.1021/jf102718t

33. Zywicki B, Reemtsma T, Jekel M (2002) Analysis of commercial vegetable tanning agents by reversed-phase liquid chromatography-electrospray ionization-tandem mass spectrometry and its application to wastewater. J Chromatogr A 970:191–200. doi: 10.1016/S0021-9673(02)00883-X

34. Vivas N, Bourgeois G, Vitry C, Glories Y, de Freitas V (1996) Determination of the Composition of Commercial Tannin Extracts by Liquid Secondary Ion Mass Spectrometry (LSIMS). J Sci Food Agric 72:309–317. doi: 10.1002/(SICI)1097-0010(199611)72:3<309::AID-JSFA658>3.0.CO;2-U

35. Abdalla S, Pizzi A, Bahabri F, Ganash A (2015) Analysis of Valonia Oak (Quercus aegylops) Acorn Tannin and Wood Adhesives Application. Bioresources 10:7165–7177

36. Mahmoud SB, Saad H, Charrier B, Pizzi A, Rode K, Ayed N, Bouhtoury FC (2015) Characterization of sumac (Rhus tripartitum) root barks tannin for a potential use in wood adhesives formulation. Wood Sci Technol 49:205–221. doi: 10.1007/s00226-014-0686-4

37. Sáyago-Ayerdi S, Moreno-Hernández C, Montalvo-González E, García-Magaña M, Lluís J, Pérez-Jiménez J (2013) Mexican ‘Ataulfo’ mango (Mangifera indica L) as a source of hydrolyzable tannins. Analysis by MALDI-TOF/TOF MS. Food Res Int 51:188–194. doi: 10.1016/j.foodres.2012.11.034

38. Saad H, Bouhtoury FC, Pizzi A, Rode K, Charrier B, Ayed N (2012) Characterization of pomegranate peels tannin extractives. Ind Crop Prod 40:239–246. doi: 10.1016/j.indcrop.2012.02.038

39. Pizzi, A. Pasch, H. Rode, K. Giovando S (2009) Polymer Structure of Commercial Hydrolyzable Tanins by Matrix-Assisted Laser Desorption/Ionization-Time-of-Flight Mass Spectriometry. J Appl Polym Sci 113:3847–3859

40. Radebe N, Rode K, Pizzi A, Giovando S, Pasch H (2013) MALDI-TOF-CID for the microstructure elucidation of polymeric hydrolysable tannins. J Appl Polym Sci 128:97–107. doi: 10.1002/app.38156

41. Buryakov IA, Krylov E V., Nazarov EG, Rasulev UK (1993) A new method of separation of multi-atomic ions by mobility at atmospheric pressure using a high-frequency amplitude-asymmetric strong electric field. Int J Mass Spectrom Ion Process 128:143–148. doi: 10.1016/0168-1176(93)87062-W

42. Kanu AB, Dwivedi P, Tam M, Matz L, Herbert H (2008) Ion mobility – mass spectrometry. J Mass Spectrom 43:1–22. doi: 10.1002/jms

43. Kolakowski BM, Mester Z (2007) Review of applications of high-field asymmetric waveform ion mobility spectrometry (FAIMS) and differential mobility spectrometry (DMS). Analyst 132:842–864. doi: 10.1039/b706039d

44. Clemmer DE, Jarrold MF (1997) Ion mobility measurement and their applications to cluster and biomolecules. J Mass Spectrom 32:577–592

45. Franceschi P, Vrhovsek U, Guella G (2011) Ion mobility mass spectrometric investigation of ellagitannins and their non-covalent aggregates. Rapid Commun Mass Spectrom 25:827–833

46. Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH (2004) Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. Rapid Commun Mass Spectrom 18:2401–2414. doi: 10.1002/rcm.1641

47. Moon MH, Myung S, Plasencia M, Hilderbrand AE, Clemmer DE (2003)

Nanoflow LC/Ion Mobility/CID/TOF for Proteomics: Analysis of a Human Urinary Proteome. J Proteome Res 2:589–597

48. Sowell A, Koeniger SL, Valentine SJ, Moon MH, Clemmer DE (2004) Nanoflow LC/IMS-MS and LC/IMS-CID/MS of Protein Mixtures. Am Soc Mass Spectrom 15:1341–1353. doi: 10.1016/j.jasms.2004.06.014

49. Glabasnia A, Hofmann T (2006) Sensory-Directed Identification of Taste-Active Ellagitannins in American (Quercus alba L.) and European Oak Wood (Quercus robur L.) and Quantitative Analysis in Bourbon Whiskey and Oak-Mutured Red Wines. J Agric Food Chem 54:3380–3390. doi: 10.1021/jf052617b

50. Ruotolo BT, Benesch JL, Sandercock AM, Hyung S, Robinson C V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes. Nat Protoc 3:1139–1152

51. Forsythe JG, Petrov AS, Walker CA, Allen SJ, Pellissier JS, Bush MF, Hud N V., Fernández FM (2015) Collision cross section calibrants for negative ion mode traveling wave ion mobility-mass spectrometry. Analyst 140:6853–6861

52. Stow SM, Causon TJ, Zheng X, Kurulugama RT, Mairinger T, May JC, Rennie EE, Baker ES, Smith RD, Mclean JA, Hann S, Fjeldsted JC (2017) An Interlaboratory Evaluation of Drift Tube Ion Mobility − Mass Spectrometry Collision Cross Section Measurements. Anal Chem 89:9048–9055. doi: 10.1021/acs.analchem.7b01729

53. Kalili KM, Cabooter D, Desmet G, Villiers A De (2012) Kinetic optimisation of the reversed phase liquid chromatographic separation of proanthocyanidins on sub-2 um and superficially porous phases. J Chromatogr A 1236:63–76. doi: 10.1016/j.chroma.2012.02.067

54. Willemse CM, Stander MA, de Villiers A (2013) Hydrophilic interaction chromatographic analysis of anthocyanins. J Chromatogr A 1319:127–140

55. Comandini P, Lerma-Garcia MJ, Simo-Alfonso EF, Toschi TG (2014) Tannin analysis of chestnut bark samples (Castanea sativa Mill.) by HPLC-DAD-MS. Food Chem 157:290–295. doi: 10.1016/j.foodchem.2014.02.003

56. Moilanen J, Salminen JP (2008) Ecologically neglected tannins and their biologically relevant activity: Chemical structures of plant ellagitannins reveal their in vitro oxidative activity at high pH. Chemoecology 18:73–83. doi: 10.1007/s00049-007-0395-7

57. Arapitsas P, Menichetti S, Vincieri FF, Romani A (2007) Hydrolyzable tannins with the hexahydroxydiphenoyl unit and the m-depsidic link: HPLC-DAD-MS identification and model synthesis. J Agric Food Chem 55:48–55

58. Yarnes CT, Boecklen WJ, Tuominen K, Salminen JP (2006) Defining phytochemical phenotypes: size and shape analysis of phenolic compounds in oaks (Fagaceae , Quercus) of the Chihuahuan Desert. Can J Bot 84:1233–1248. doi: 10.1139/B06-076

59. Moilanen J, Koskinen P, Salminen JP (2015) Distribution and content of ellagitannins in finnish plant species. Phytochemistry 116:188–197. doi: 10.1016/j.phytochem.2015.03.002

60. Kuhnert N, Yassin GH, Jaiswal R, Matei MF, Grün CH (2015) Differentiation of prototropic ions in regioisomeric caffeoyl quinic acids by electrospray ion mobility mass spectrometry. Rapid Commun Mass Spectrom 29:675–680

61. Shvartsburg AA, Jarrold MF (1996) An exact hard-spheres scattering model for the mobilities of polyatomic ions. Chem Phys Lett 261:86–91

62. Mesleh MF, Hunter JM, Shvartsburg AA, Schatz GC, Jarrold MF (1996) Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential. J Phys Chem 100:16082–16086

63. Pringle SD, Giles K, Wildgoose JL, Williams JP, Slade SE, Thalassinos K, Bateman RH, Bowers MT, Scrivens JH (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. Int J Mass Spectrom 261:1–12. doi: 10.1016/j.ijms.2006.07.021

64. Domon B, Costello CE (1988) A Systematic Nomenclature for Carbohydrate Fragmentations in FAB-MS / MS Spectra of Glycoconjugates. 397–409

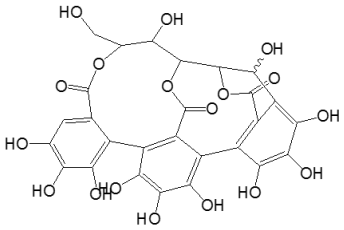
65. Ricci A, Lagel MC, Parpinello GP, Pizzi A, Kilmartin PA, Versari A (2016) Spectroscopy analysis of phenolic and sugar patterns in a food grade chestnut tannin. Food Chem 203:425–429

66. Lagel MC, Pizzi A, Giovando S (2014) Matrix-Assisted Laser Desorption-Ionization Time of Flight (MALDI-TOF) Mass Spectrometry of Phenol-Formaldehyde-Chestnut Tannin Resins. J Renew Mater 2:207–219

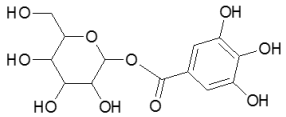
Chapter 3

Supporting information for: Comprehensive analysis of hydrolysable tannins by reversed phase and hydrophilic interaction chromatography coupled to ion mobility and high resolution mass spectrometry, Part 1: Chestnut tannins

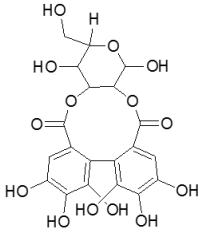
Compound 1: β-Vescalin/α-Castalin (2)



Compound 2: Monogalloyl glucose (13)



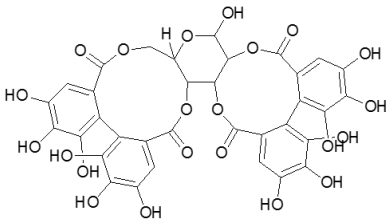
Compound 3: HHDP-glucose (4)*



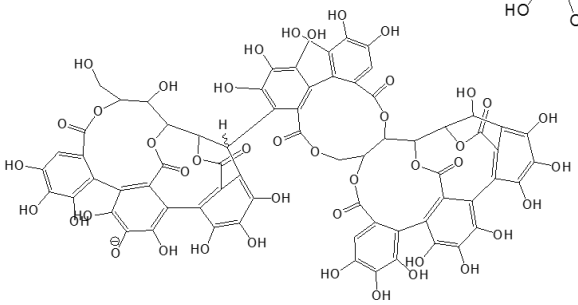
Compound 4: Pedunculagin-CO_2
No structure could be proposed based on the available data. (1)*

Compound 5: Dihydrated vescalagin and castalagin (positions of hydration unknown) (2)*

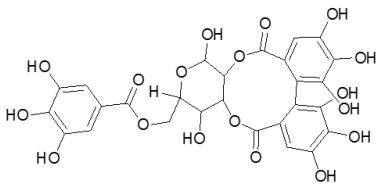
Compound 6: Pedunculagin (5)*



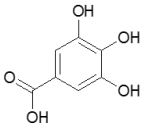
Compounds 7[a+b]: Castalagin-castalin/vescalin dimer (2)*



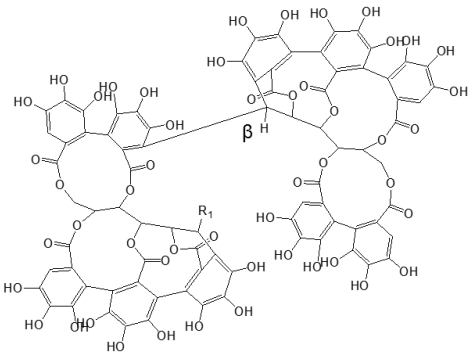
Compound 8: Galloyl-HHDP-glucose (11)*



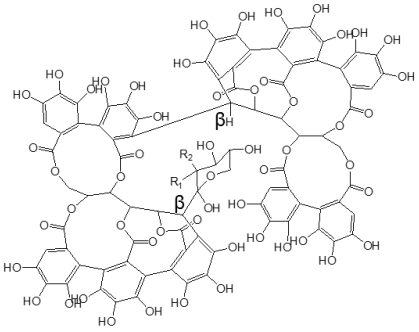
Compound 9: Gallic acid (1)



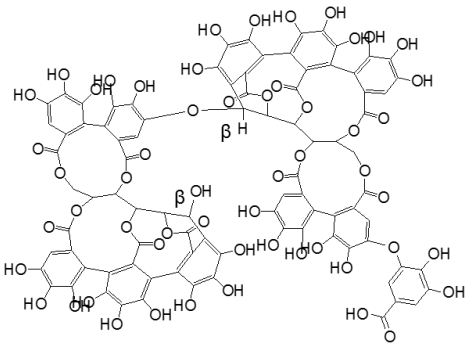
Compound 10[a]:R_1 = β-OH, Roburin A (1)* (2)[#]
Compound 10[b]:R_1 = α-OH, Roburin D (1)* (2)[#]



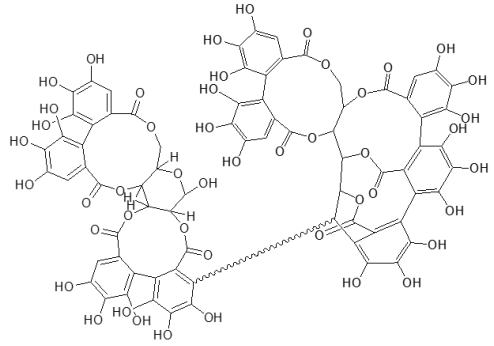
Compound 11[a]: Roburin B : R_1 = OH, R_2 = H; Xylose (1)* (2)[#]
Compound 11[b] :Roburin C: R_1 = H, R_2 = OH; Lyxose (1)* (2)[#]



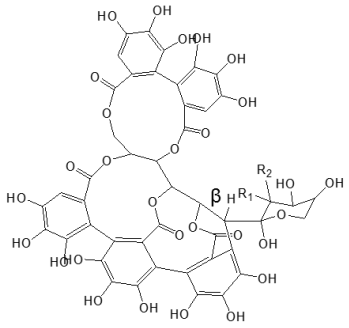
Compound 12: C-galloylated vescalagin dimer (1)* (2)[#]



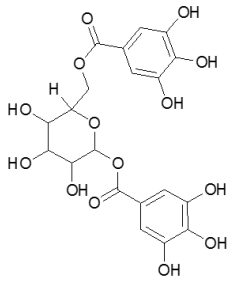
Compound 13: Vescalagin/castalagin-pedunculagin dimer (4)*



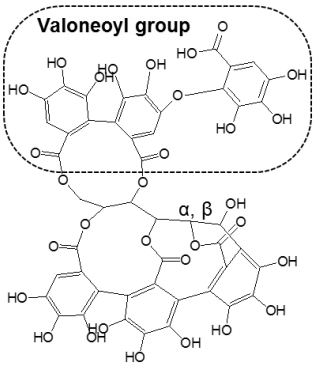
Compound 14[a]: Grandinin, R_1 = H, R_2 = OH; Lyxose (1)
Compound 14[b]: Roburin E: R_1 = OH, R_2 = H; Xylose (1)



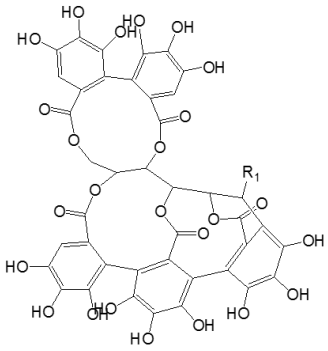
Compound 15 : Digalloyl glucose (11)



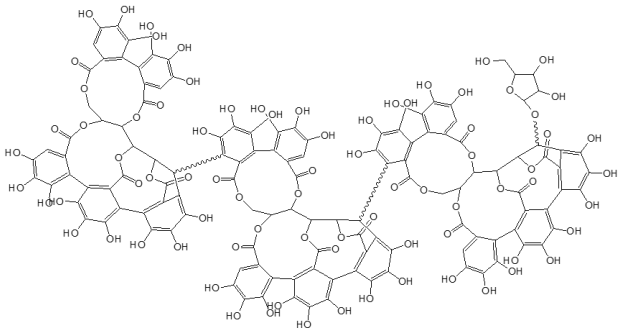
Compound 16[a]: β-Vescavaloninic acid (1)*
Compound 16[b]: α-Castavaloninic acid (1)*



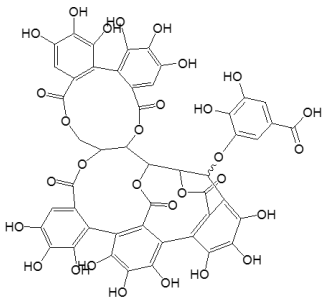
Compound 17[a]: Vescalagin, R_1 = β-OH (1)
Compound 17[b]: Castalagin, R_1 = α-OH (1)



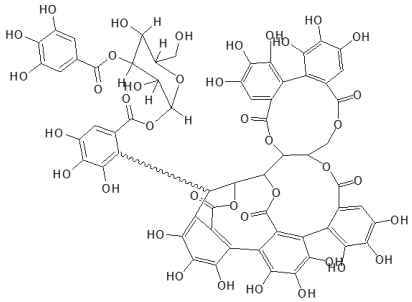
Compound 18: Glycosylated Vescalagin/Castalagin trimer (2)*(2)[#]



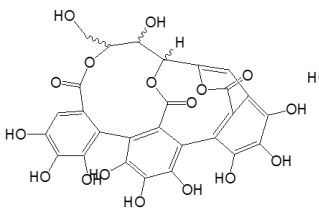
Compound 19: Galloylated vescalagin/castalagin (1)*



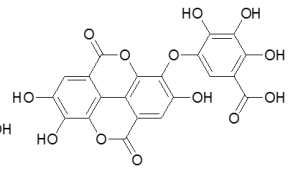
Compound 20: Vescalagin/Castalagin digalloyl glucose dimer (16)*



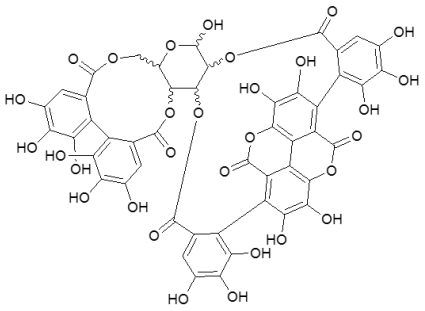
Compound 21: Vescalene (3)*



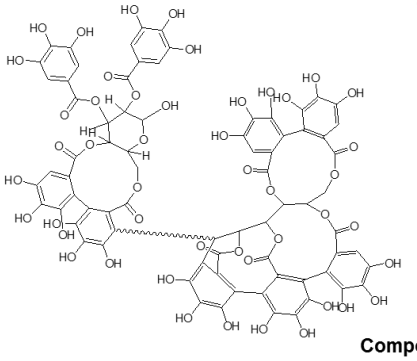
Compound 22: Valoneic acid dilactone (2)



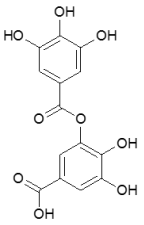
Compound 23: Punicalagin (1)*



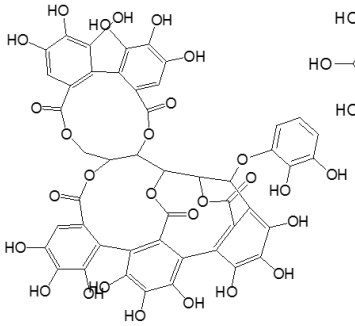
Compound 24: Vescalagin/castalagin-tellimagrandin I dimer (3)*



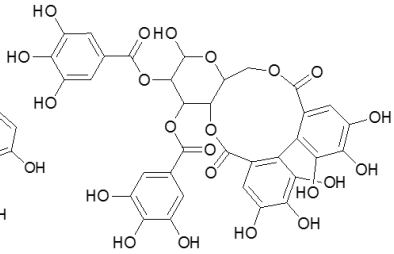
Compound 25: Digallic acid (2)*



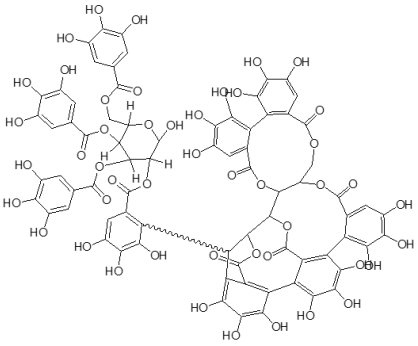
Compound 26: 1-O-trihydroxybenzene castalagin (1)*



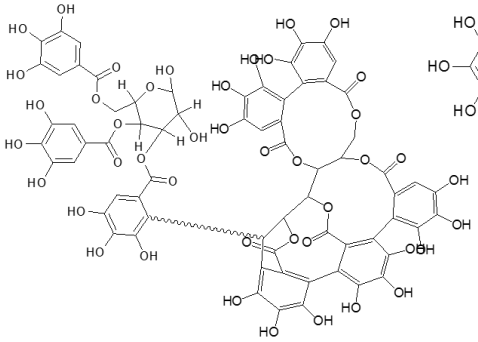
Compound 27: Tellimagrandin I (4)



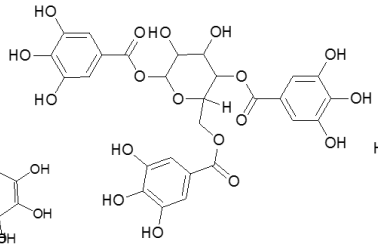
Compound 28: Vescalagin/Castalagin tetragalloyl glucose dimer (7)*



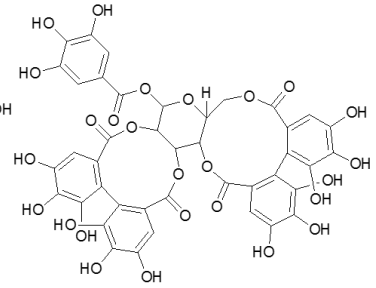
Compound 29: Vescalagin/Castalagin-trigalloyl glucose dimer (6)*



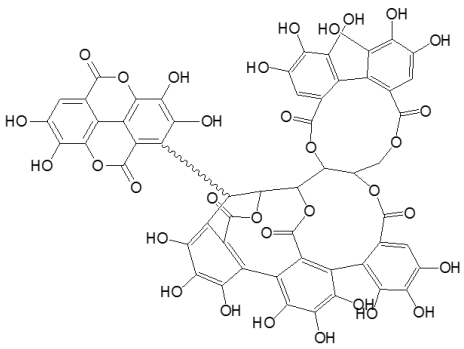
Compound 30: Trigalloyl glucose (12)



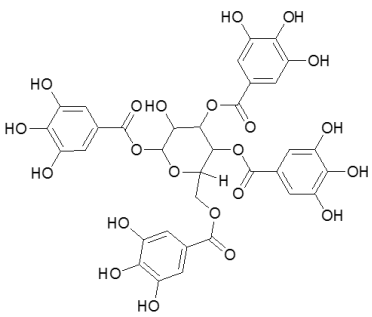
Compound 31: Casuarictin/potentillin (1)*



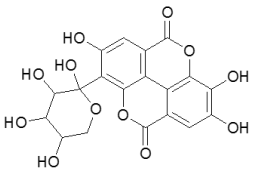
Compound 32: Vescalagin/castalagin-ellagic acid dimer (1)*



Compound 33: Tetragalloyl glucose (4)



Compound 34: Ellagic acid pentoside (1)



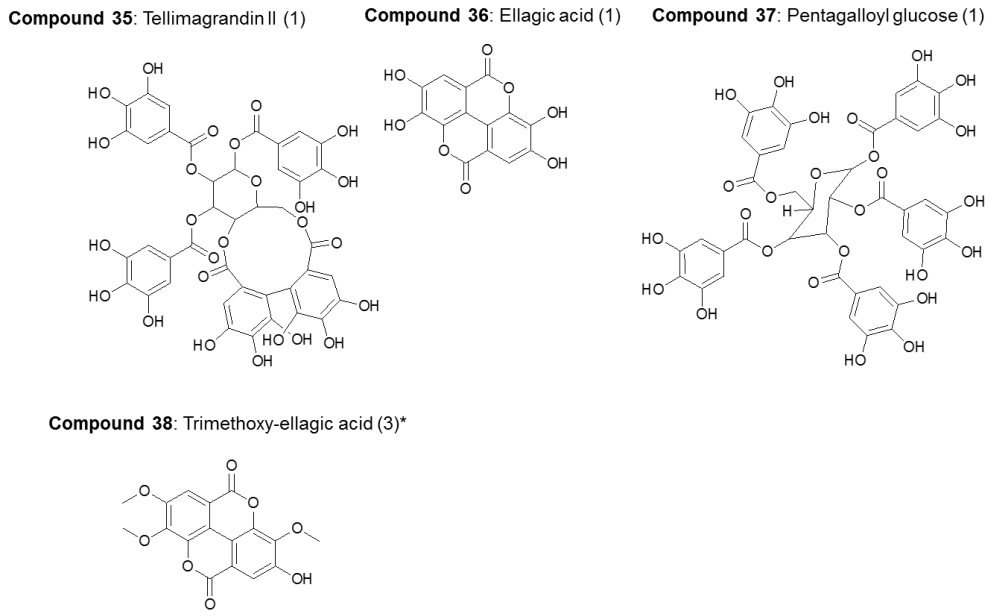


Figure S1: The chemical structures of the compounds identified in chestnut in the present work. Compounds marked with * are tentatively identified in chestnut for the first time. The number between brackets () indicates the number of isomers resolved. ()# Indicate the number of isomeric species resolved using ion mobility

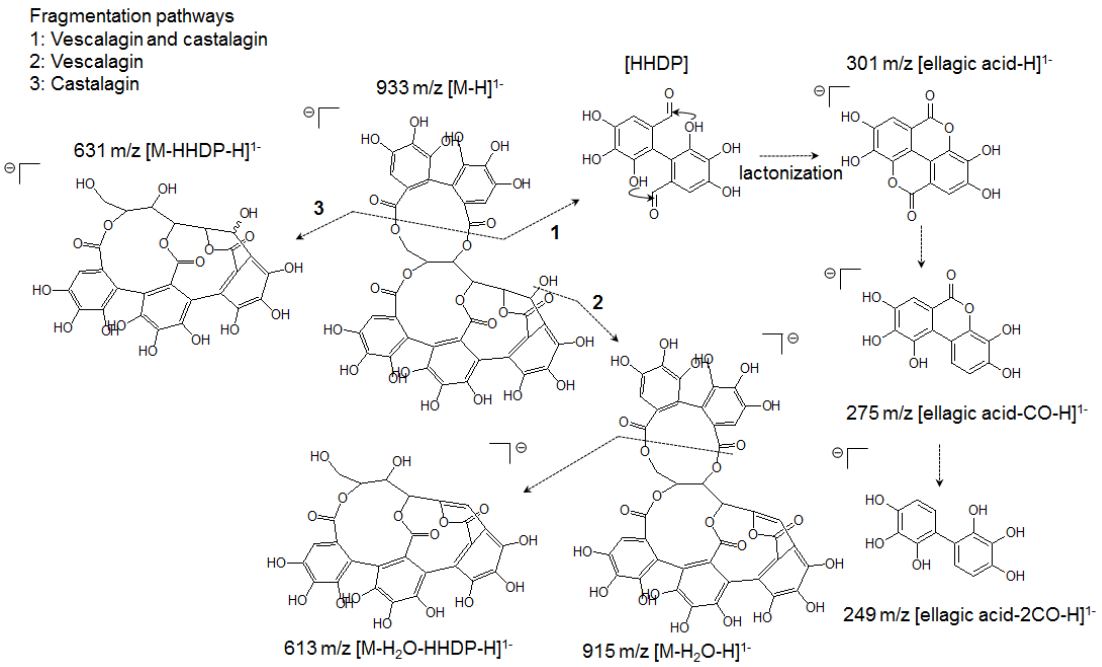


Figure S2: Illustration of the fragmentation behaviour of ellagitannins containing an HHDP unit, illustrated for vescalagin (**17[a]**) and castalagin (**17[b]**).

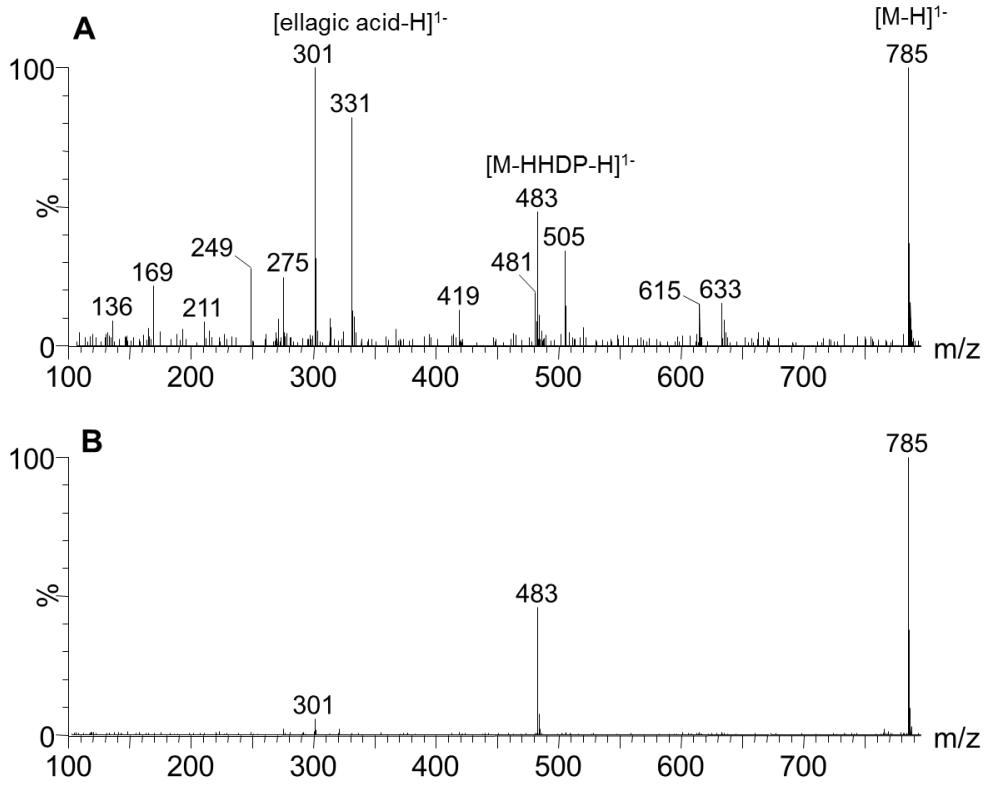


Figure S3: MS^E spectra of tellimagrandin I (**27**) obtained under HILIC **(A)** and RP-LC **(B)** conditions.

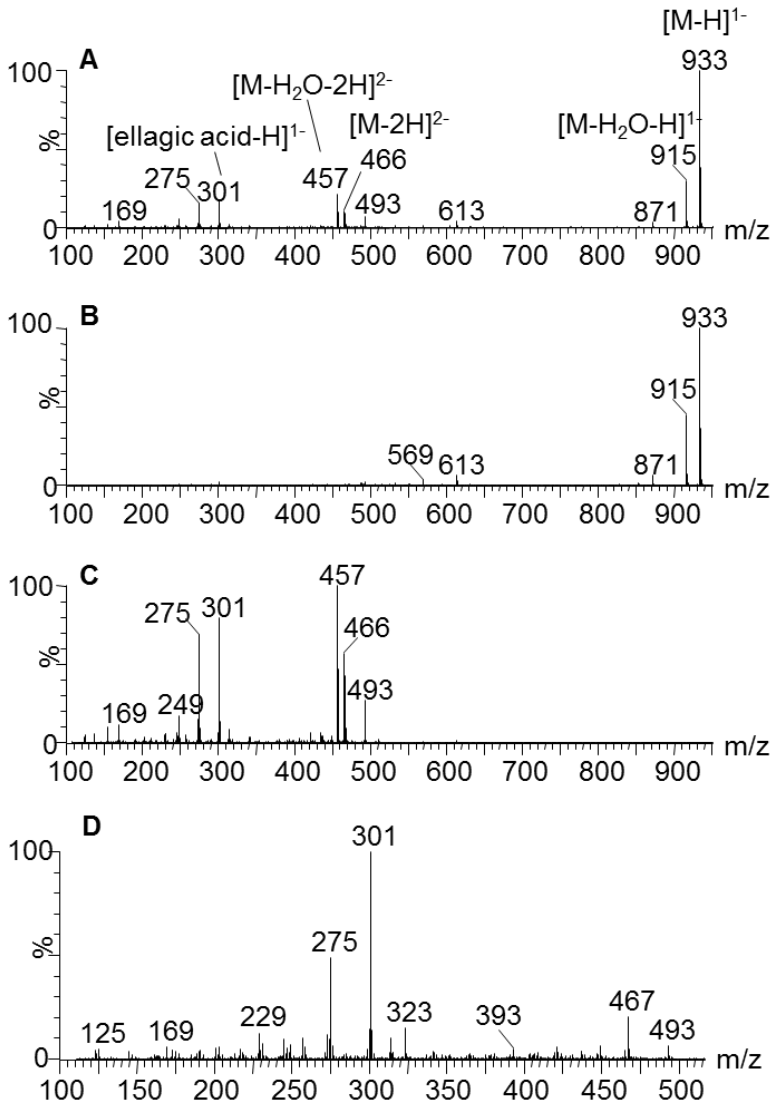


Figure S4: Comparison of MS^E spectra obtained for vescalagin (**17ª**) under different conditions. **A**: complete MS^E spectrum obtained for the RP-LC chromatographic peak at retention time 4.12 min, **B**: MS^E spectrum obtained for the singly charged ion (933 m/z, [M-H]) ion at arrival time = 6.62 ms in RP-LC, **C**: MS^E spectrum for the doubly charged ion (466 m/z, $[M-2H]^{2-}$) obtained at a arrival time of 1.93 ms in RP-LC. **D**: MS^E spectrum for the doubly charged ion (466 m/z $[M-2H]^{2-}$) at 1.93 ms in HILIC.

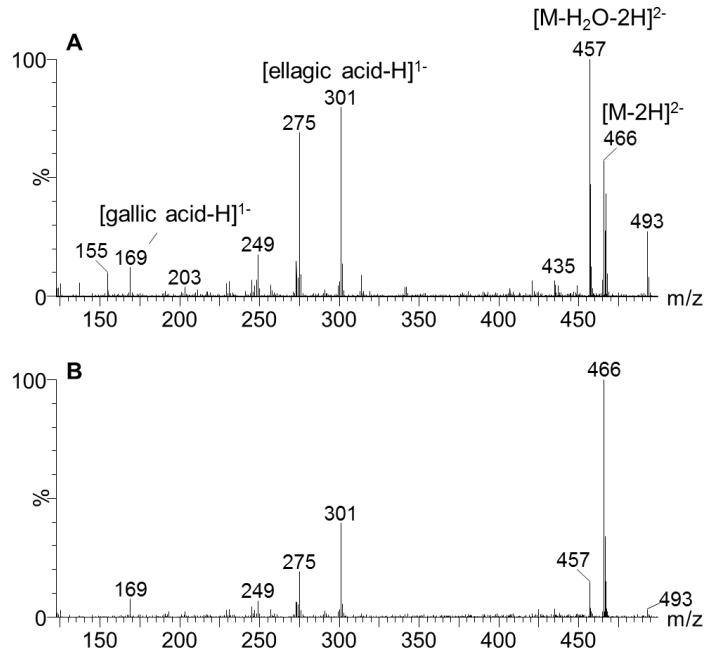


Figure S5: Comparison of the MS^E spectra of the doubly charged molecular ions (466 m/z, $[M-2H]^{2-}$) of vescalagin (**A**) and castalagin (**B**). Presented spectra were obtained from RP-LC analysis.

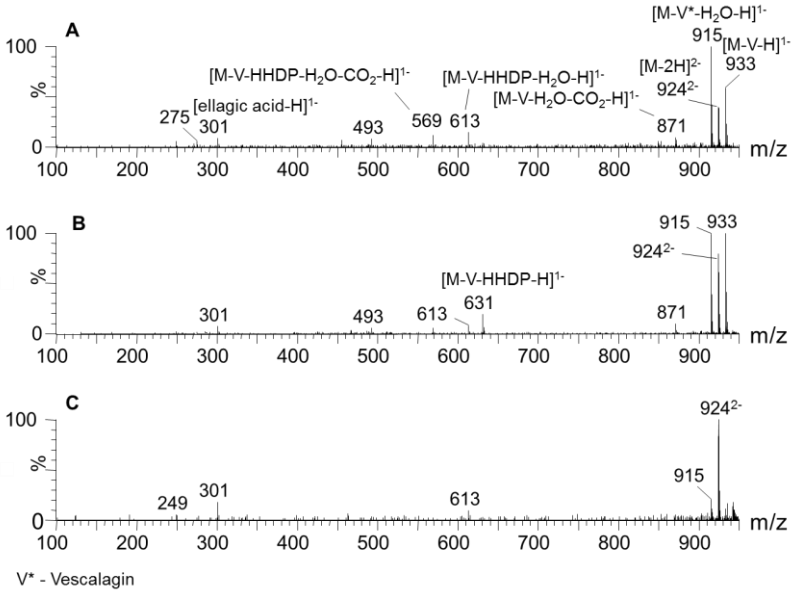


Figure S6: MS^E spectra of roburin A (**10^a**) (**A**, arrival times 3.52 and 4.35ms) and D (**10^b**) (**B**, arrival times 3.52 and 4.35 ms) and a tentatively identified dimeric species (**10^c**) (**C**, arrival time 4.00 ms) obtained in RP-LC.

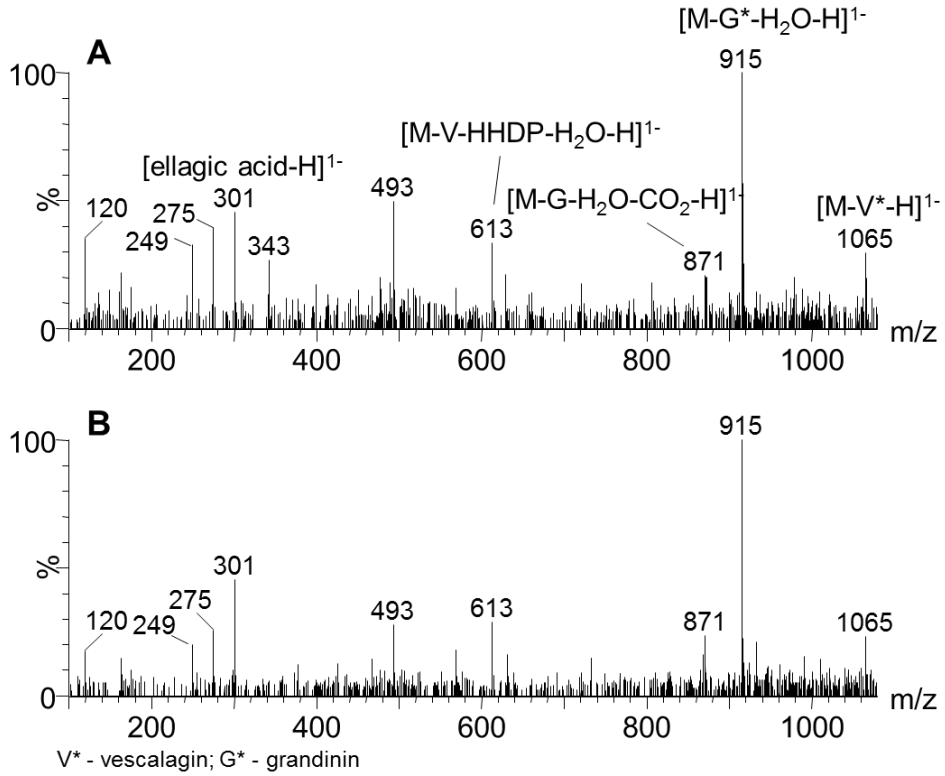


Figure S7: MS^E spectra of (**A**) roburin B (**11ª**) and (**B**) roburin C (**11ᵇ**) obtained by RP-LC.

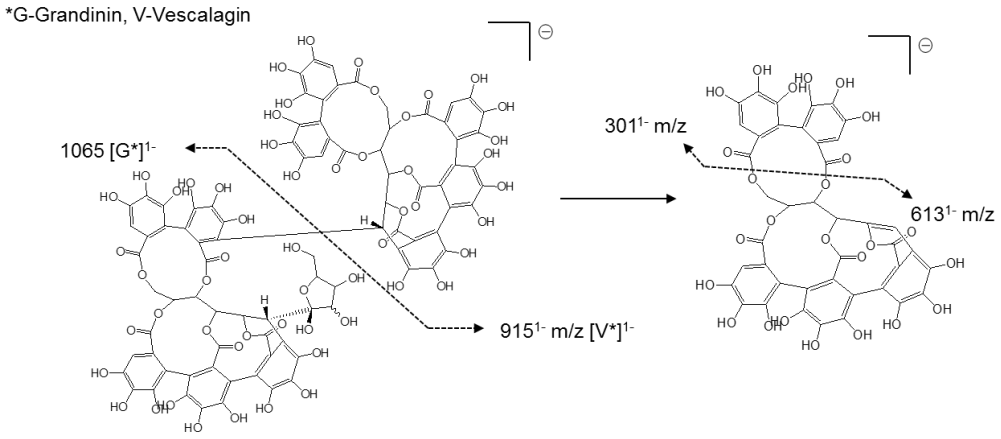


Figure S8: Proposed fragmentation of roburin B (**11ª**) and roburin C (**11ᵇ**).

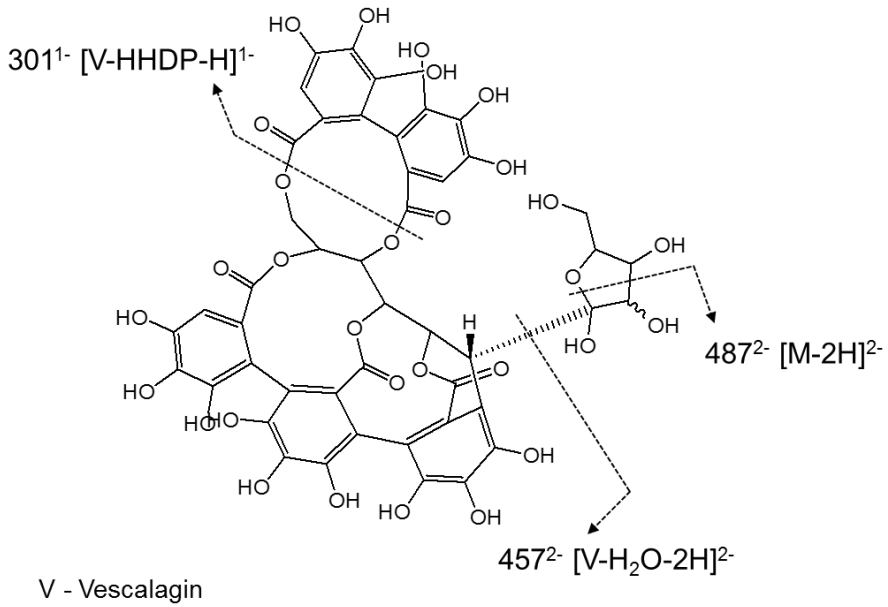


Figure S9: Proposed fragmentation for grandinin (**14[a]**) and roburin E (**14[b]**).

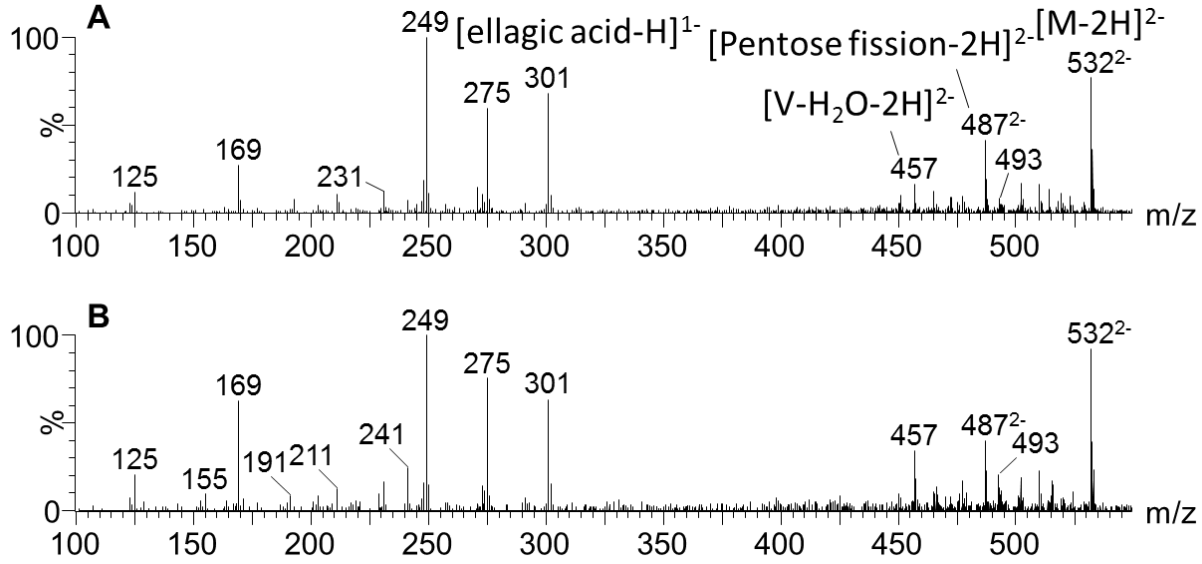


Figure S10: MS[E] spectra of the doubly charged vescalagin-pentoside isomers (532 m/z) grandinin (**14[a]**) and roburin E (**14[b]**) at retention times 2.98 (**A**) and 4.10 (**B**), respectively, in RP-LC.

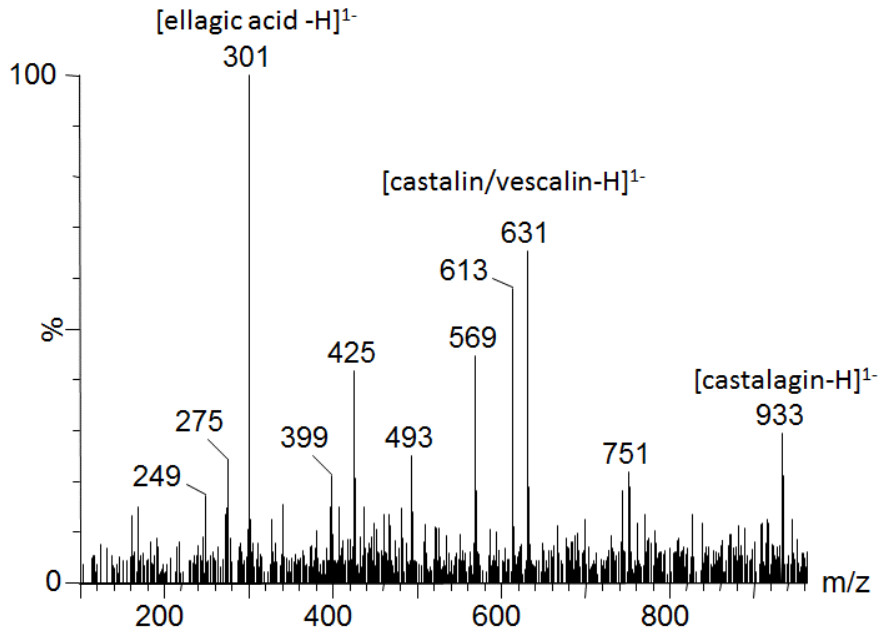


Figure S11: MS^E spectrum obtained for the doubly charged molecular ion (773 m/z, $[M-2H]^{2-}$) of a tentatively identified castalagin-castalin/vescalin dimer (**7**) obtained by RP-LC.

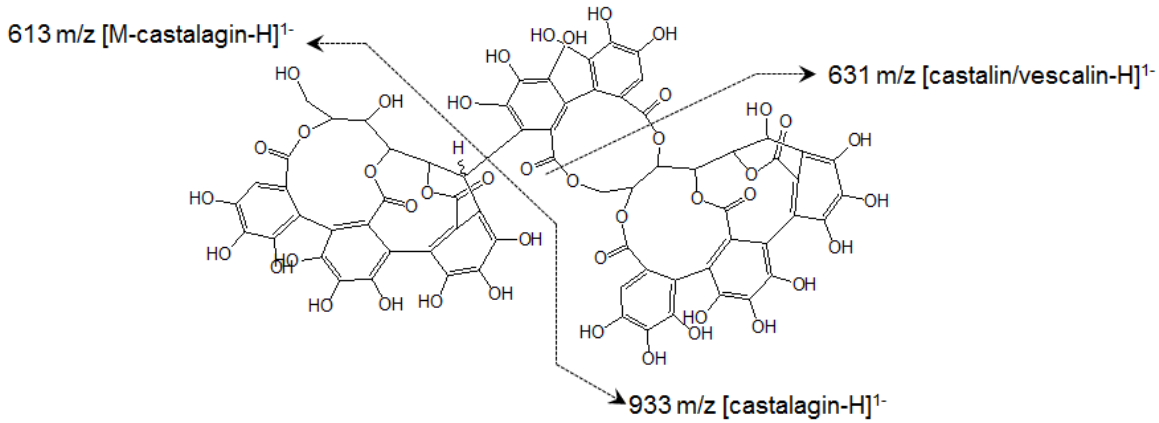


Figure S12: Proposed fragmentation scheme of a tentatively identified castalagin-castalin/vescalin dimer (**7**).

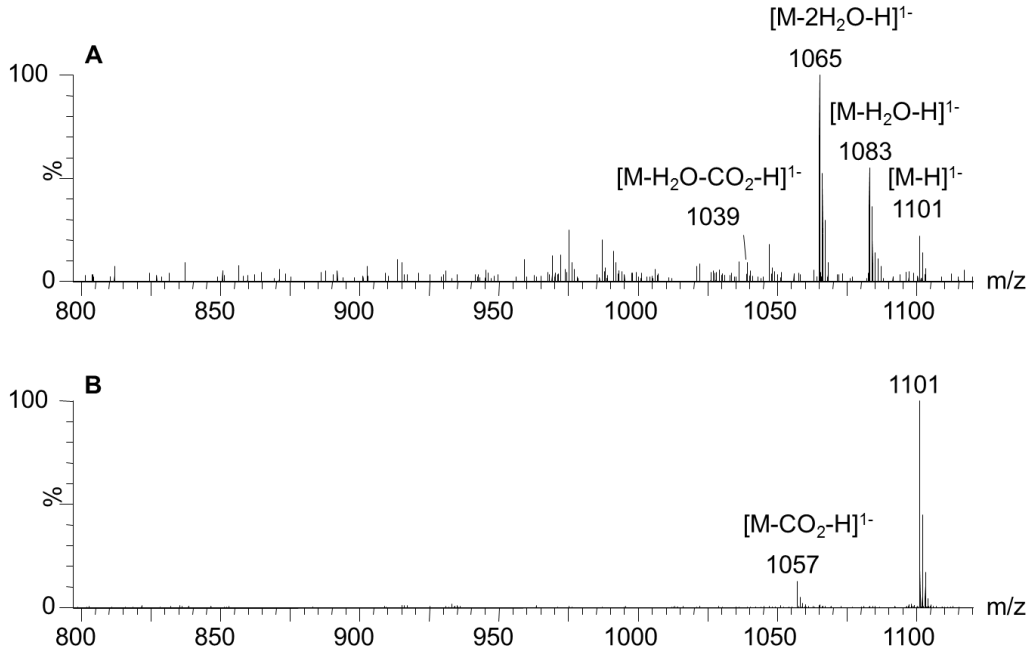


Figure S13: MS^E spectra of (**A**) vescavaloninic acid (**16^a**) and (**B**) castavaloninic acid (**16^b**) obtained by RP-LC.

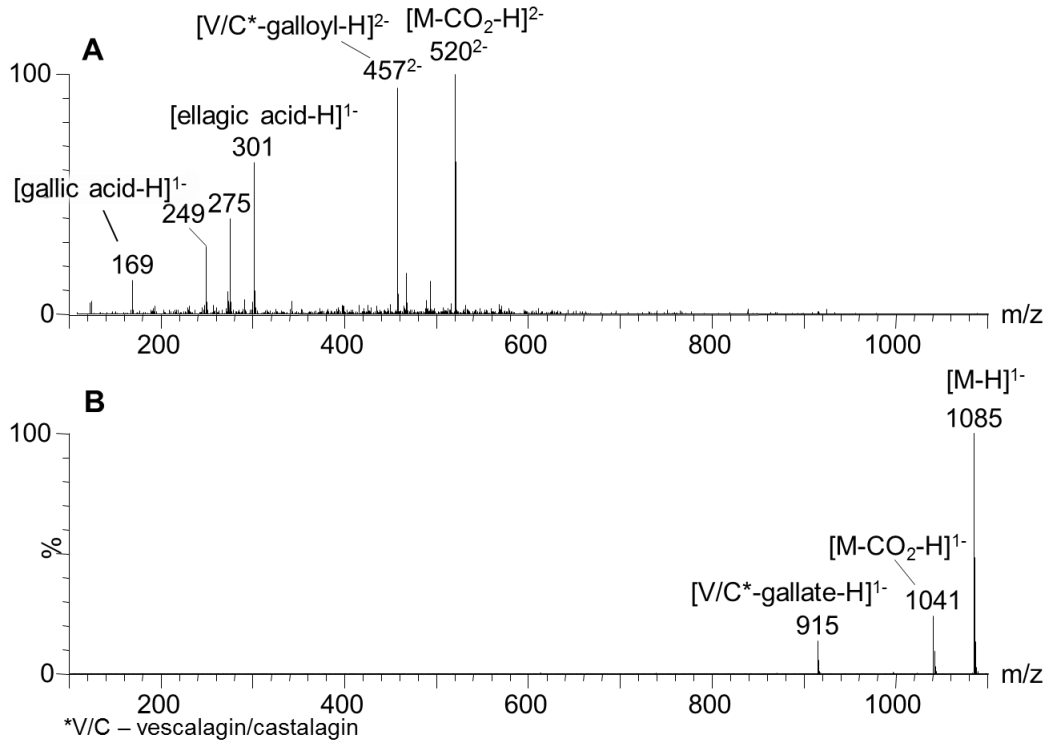


Figure S14: MS^E spectra of the doubly (**A**) and singly (**B**) charged molecular ions of galloylated vescalagin/castalagin (**19**). V = vescalagin, C = castalagin.

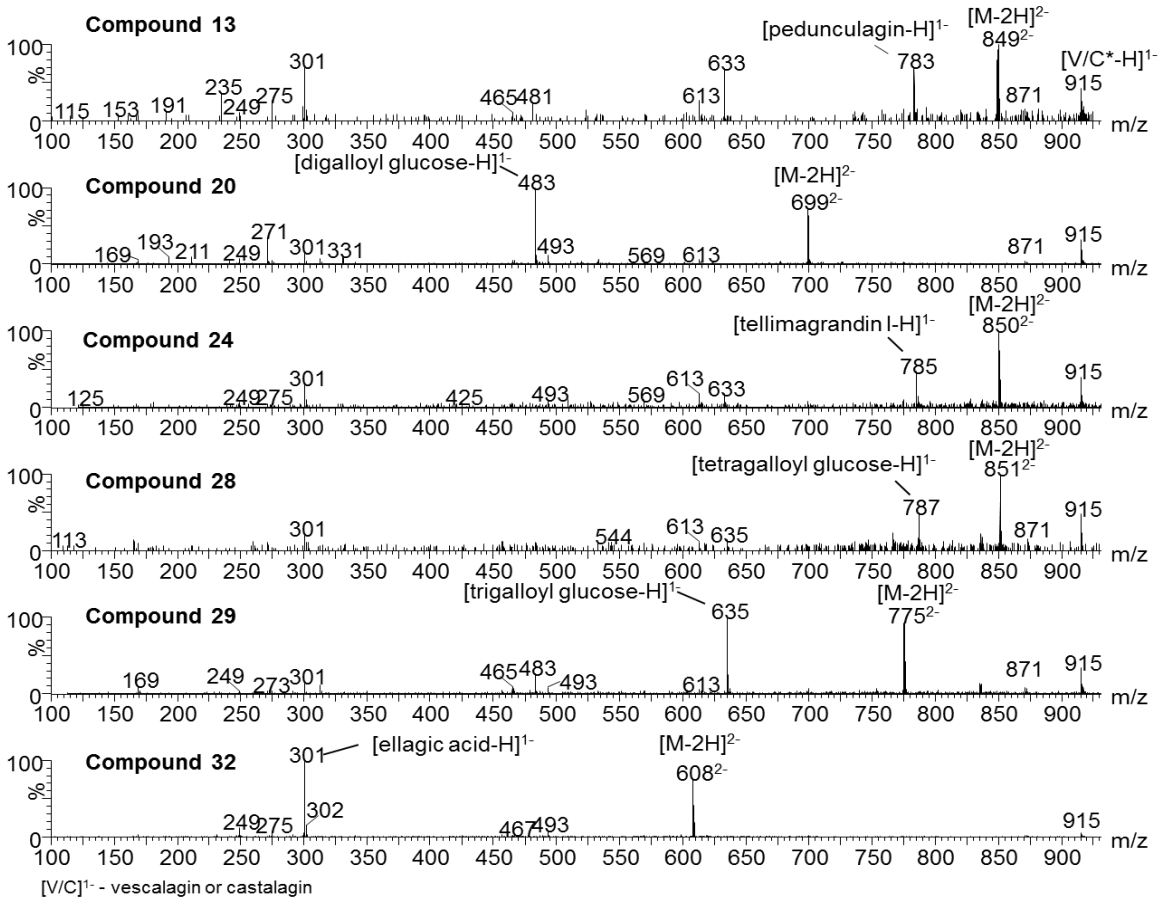


Figure S15: Examples of MS^E spectra obtained for the doubly charged molecular ions of compounds **13, 20, 24, 28, 29** and **32**. Spectra obtained from the RP-LC analysis of chestnut. For interpretation of observed fragments, refer to **Figure S16**. V = vescalagin, C = castalagin.

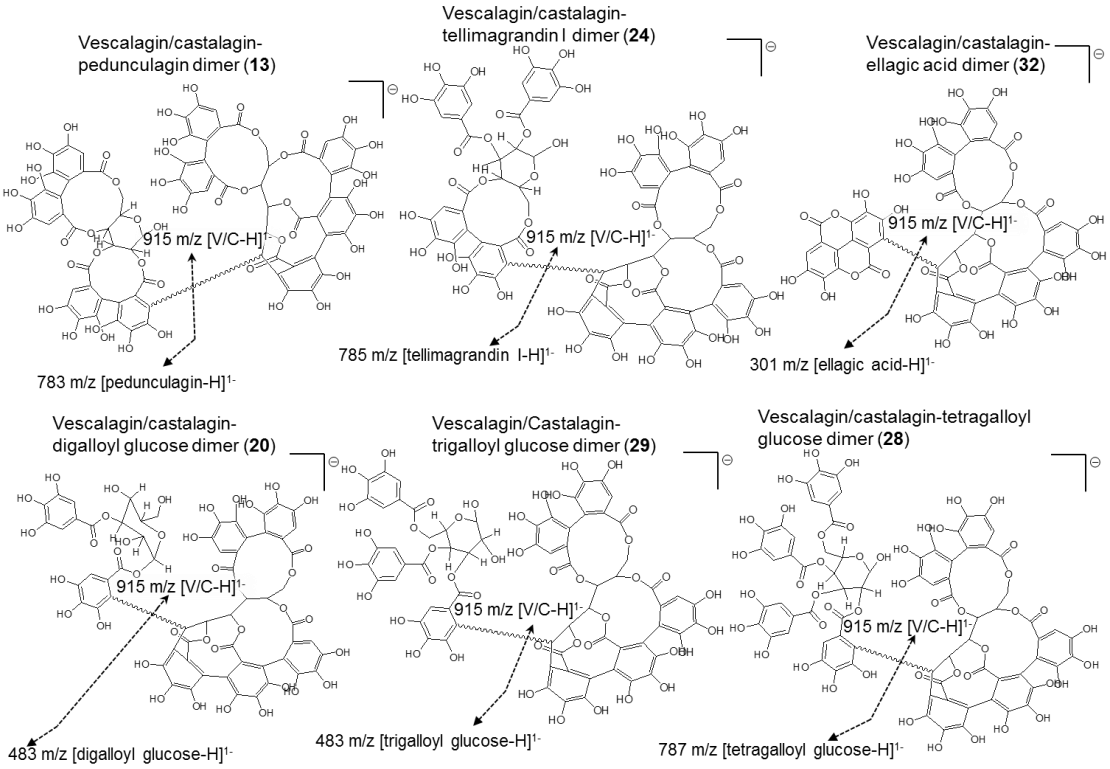


Figure S16: Proposed fragmentation schemes for the dimeric castalagin/vescalagin-closed glucose ellagitannin derivatives **13**, **20**, **24**, **28**, **29** and **32**.

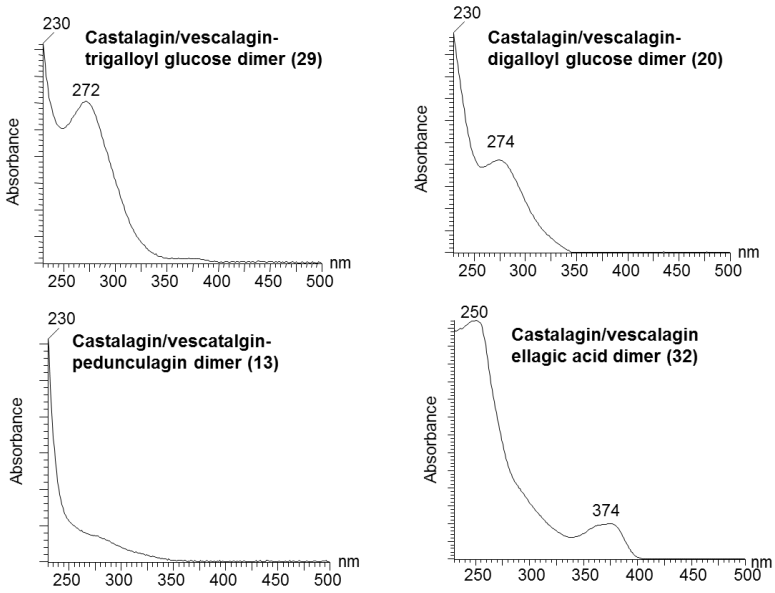


Figure S17: Examples of on-line UV spectra recorded for selected isomers of compounds **13**, **20**, **29** and **32**.

Table S1: Summary of non-covalent cluster ions detected for vescalagin (**17**[a]) and castalagin (**17**[b]) by RP-LC- and HILIC-ESI-IM-HR-MS.

RP-LC	HILIC	t_R RP-LC	t_R HILIC (min)	Exp. m/z	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$
$[2V-2H]^{2-}$	$[2C-2H]^{2-}$	4.12	21.54	933.0652	1.29	3.11	392
$[2C-2H]^{2-}$	$[2V-2H]^{2-}$	5.63	17.28	933.0665	2.68		
$[3V-2H]^{2-}$	$[3C-2H]^{2-}$	4.12	21.54	1400.1127	4.14	5.11	503
$[3C-2H]^{2-}$	$[3V-2H]^{2-}$	5.63	17.28	1400.1133	4.57		
$[5V-3H]^{3-}$	nd	4.12	nd	1555.7694	-5.27	4.07, 4.28	673
$[5C-3H]^{3-}$	nd	5.63	nd	1555.7780	0.26	4.07, 4.28	690
$[V+G/R-2H]^{2-}$	nd	4.12	nd	999.0820	-2.50	3.38	409
nd	$[2V-3H+Na^+]^{2-}$	nd	17.28	944.0591	4.98	3.17, 3.66	396
nd	$[2C-3H+Na^+]^{2-}$	nd	21.54	944.0568	2.54	3.17, 3.66	426

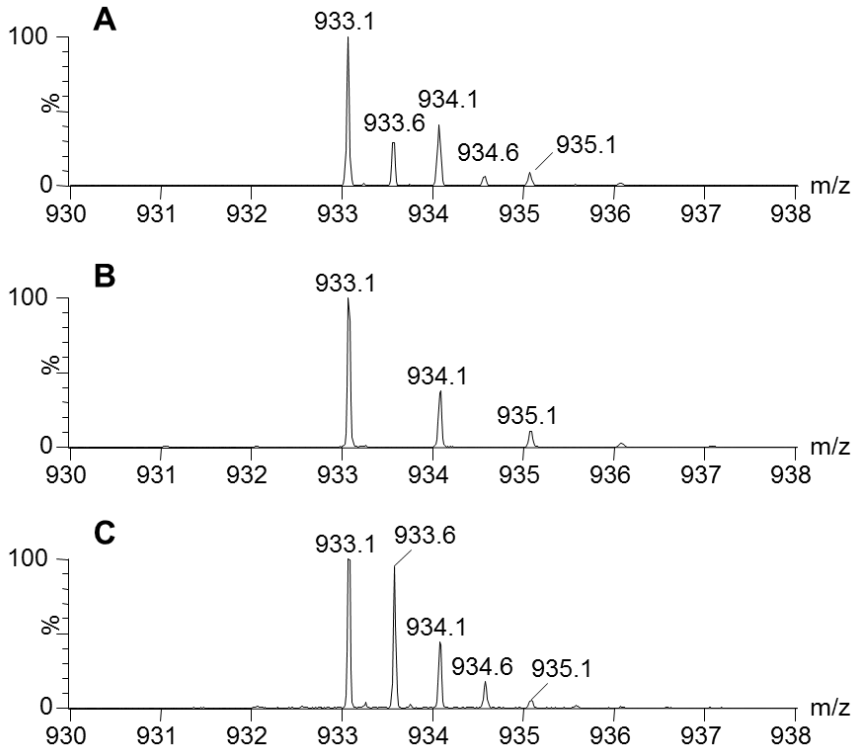


Figure S18: Isotope distributions of the singly charged molecular ion region in the low energy mass spectra obtained for vescalagin (**17**[a]) (**A**) from the chromatographic trace, (**B**) at arrival time 6.62 ms ($[V-H]^{1-}$), and (**C**) at arrival time 3.38 ms ($[2V-2H]^{2-}$).

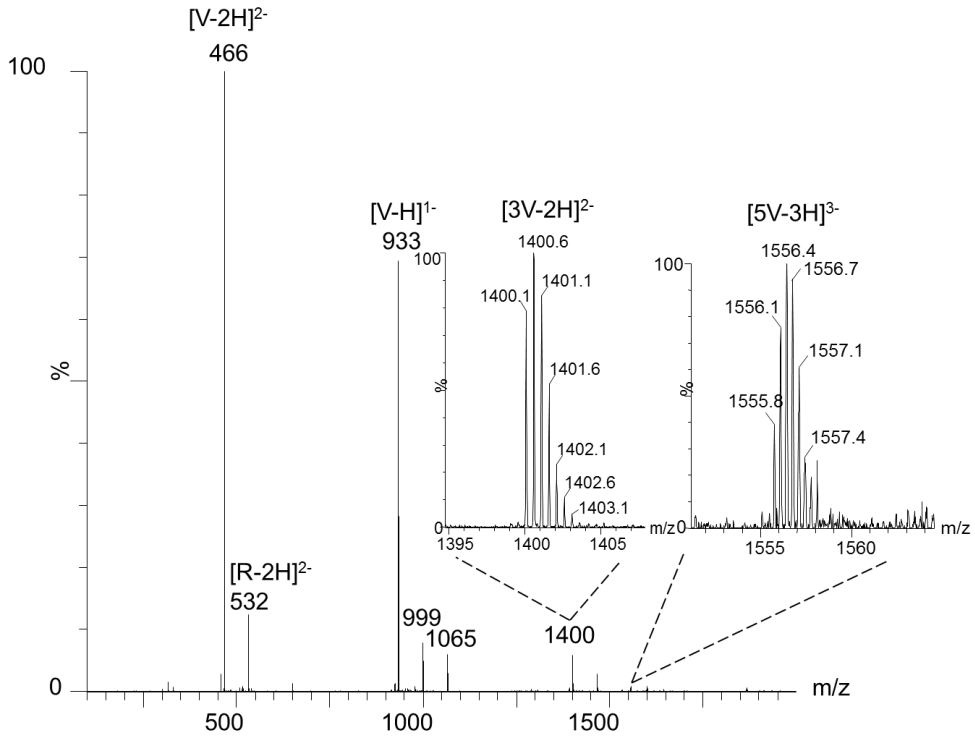


Figure S19: Low collision energy mass spectrum of vescalagin (**17[a]**) obtained from the RP-LC chromatographic trace showing the presence of the doubly charged trimeric ($[3V-2H]^{2-}$) and triply charged pentameric $[5V-3H]^{3-}$ cluster ions.

Table S2: Summary of the multiply charged heteroisotopic clusters detected for vescalagin.

Clusters	Calculated mass	Experimental mass (vescalagin)	Mass error (ppm)
$[4,^{13}C(1)]^{8-}$	466.1532	466.1539	1.48
$[2,^{13}C(1)]^{4-}$	466.2786	466.2801	3.15
$[4,^{13}C(3)]^{8-}$	466.6549	466.6597	10.31
$[2,^{13}C(3)]^{4-}$	466.7803	466.7780	-4.94
$[4,^{13}C(5)]^{8-}$	467.1566	467.1637	15.27

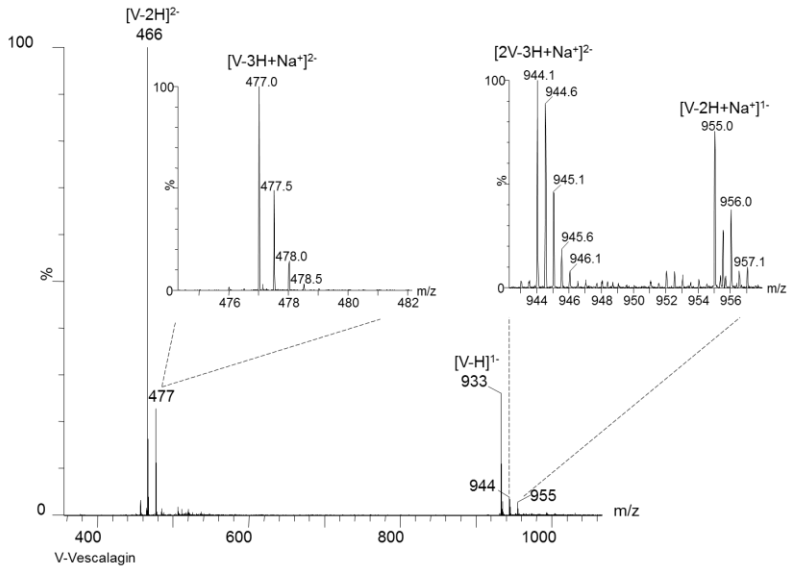


Figure S20: Example of the low collision energy spectrum obtained for vescalagin (**17[a]**) from the HILIC chromatographic trace illustrating the ionic species detected (**Table S1**).

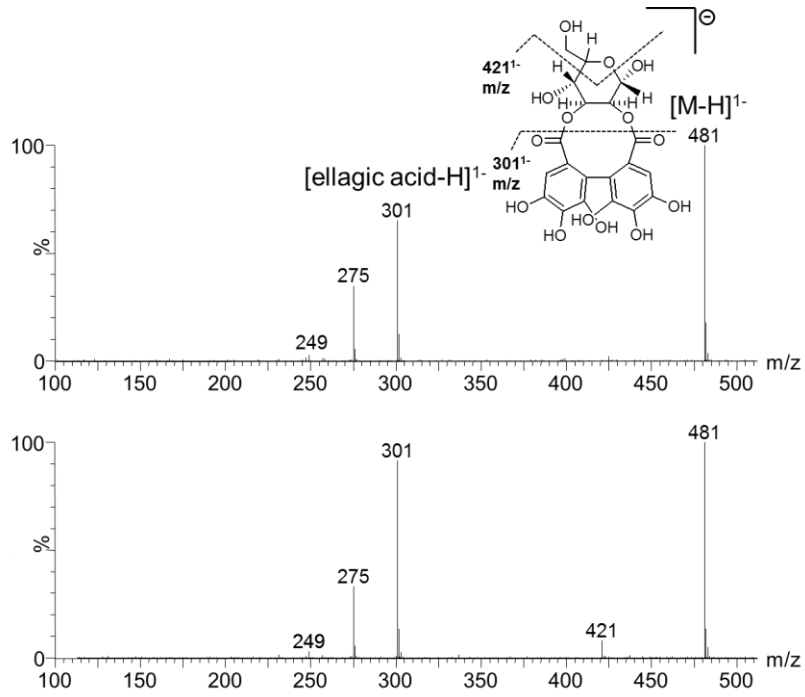


Figure S21: Example MS[E] spectra obtained in RP-LC of HHDP-glucose isomers (**A**) **3[a]** and (**B**) **3[b]**. (Identical spectra were obtained for isomers **3[a]** and **3[d]**, as well as for isomers **3[b]** and **3[c]** (**Table 2**).

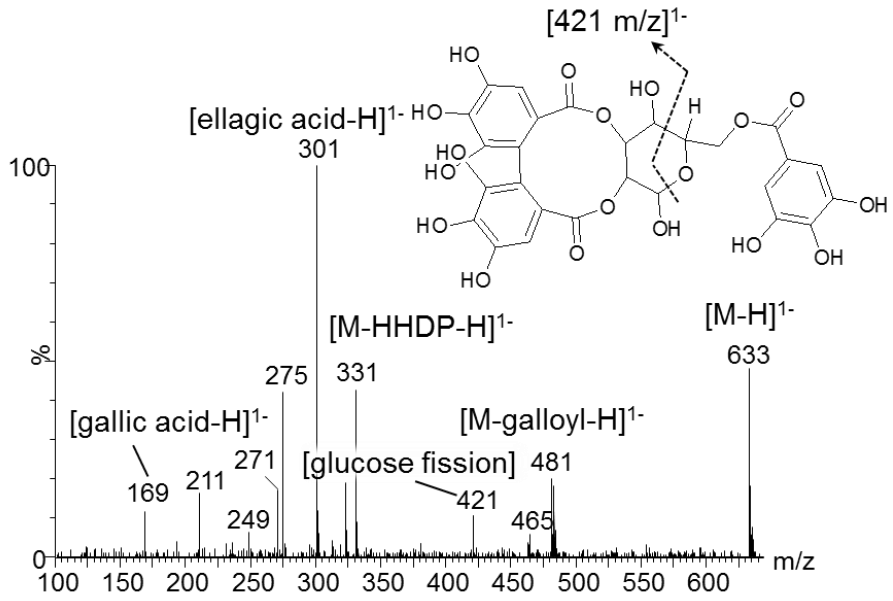


Figure S22: MS^E spectrum of a galloyl-HHDP-glucose isomer eluting at 9.11 min in HILIC. The 421 m/z fragment ion is proposed to be due to intra-molecular cleavage of the sugar moiety, and can therefore be used to confirm the position of galloylation as being C-6.

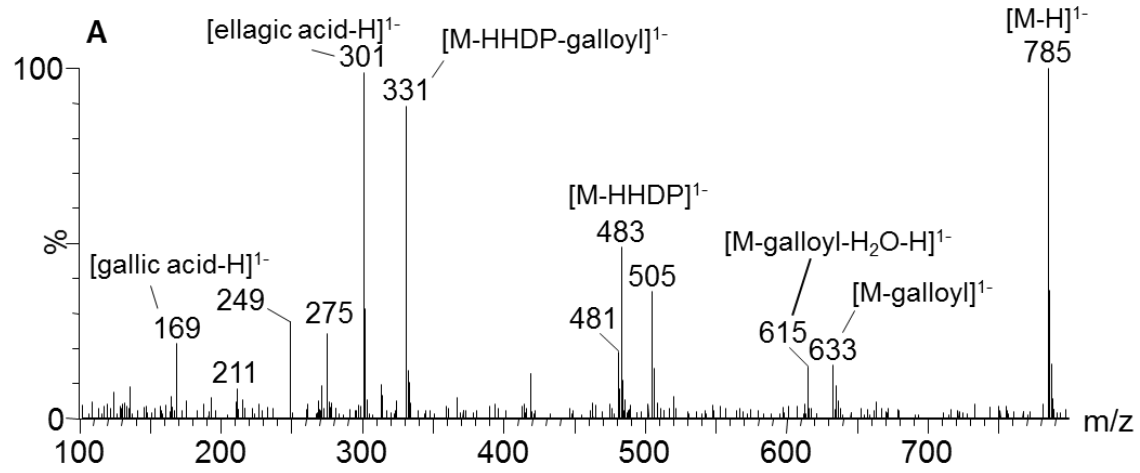


Figure S23: MS^E spectrum obtained for a tellimagrandin I (**27**) isomer eluting at 8.97 min in the HILIC analysis.

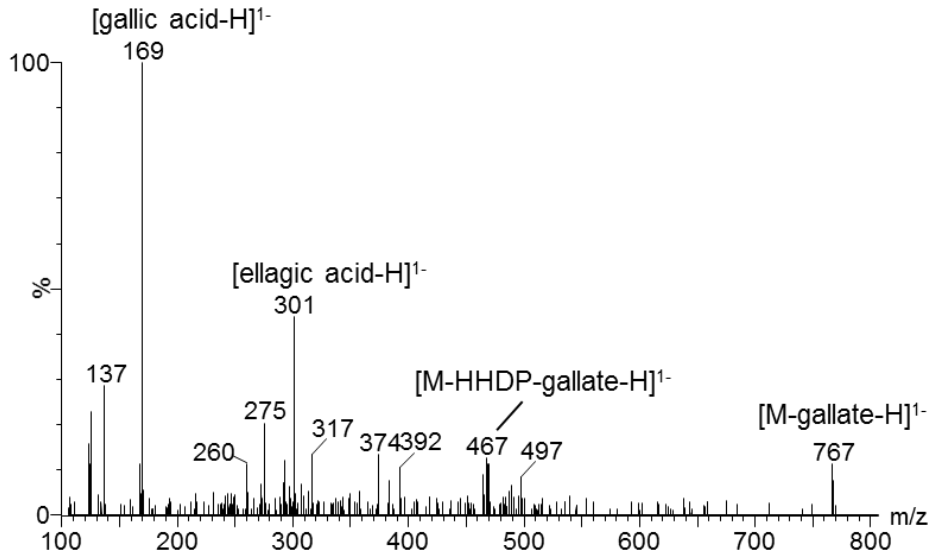


Figure S24: MS^E spectrum obtained for a tellimagrandin II (**35**) isomer eluting at 11.23 min in the RP-LC analysis.

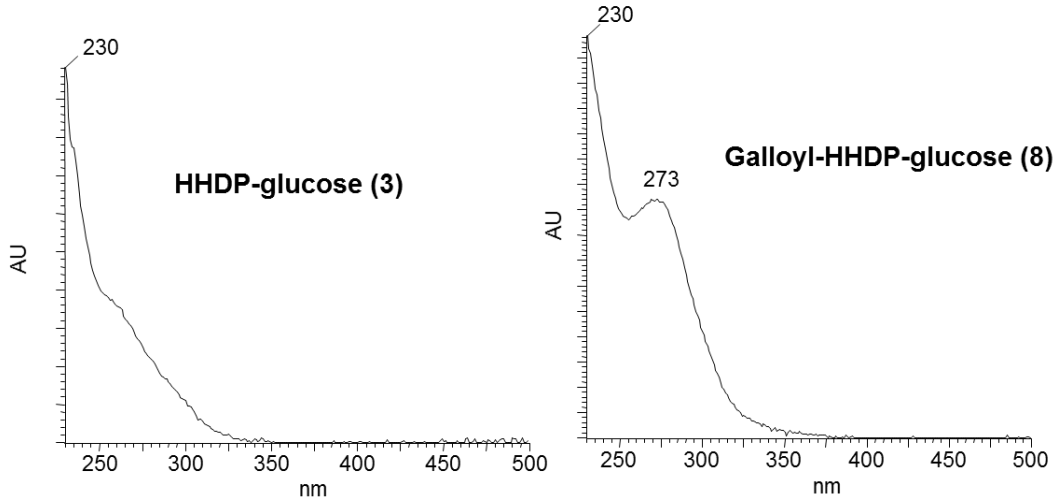


Figure S25: Typical examples of UV spectra obtained for closed glucose ellagitannins HHDP-glucose (**3**) and galloyl-HHDP-glucose (**8**) illustrating the effect of galloylation on relative absorbance at 273 nm.

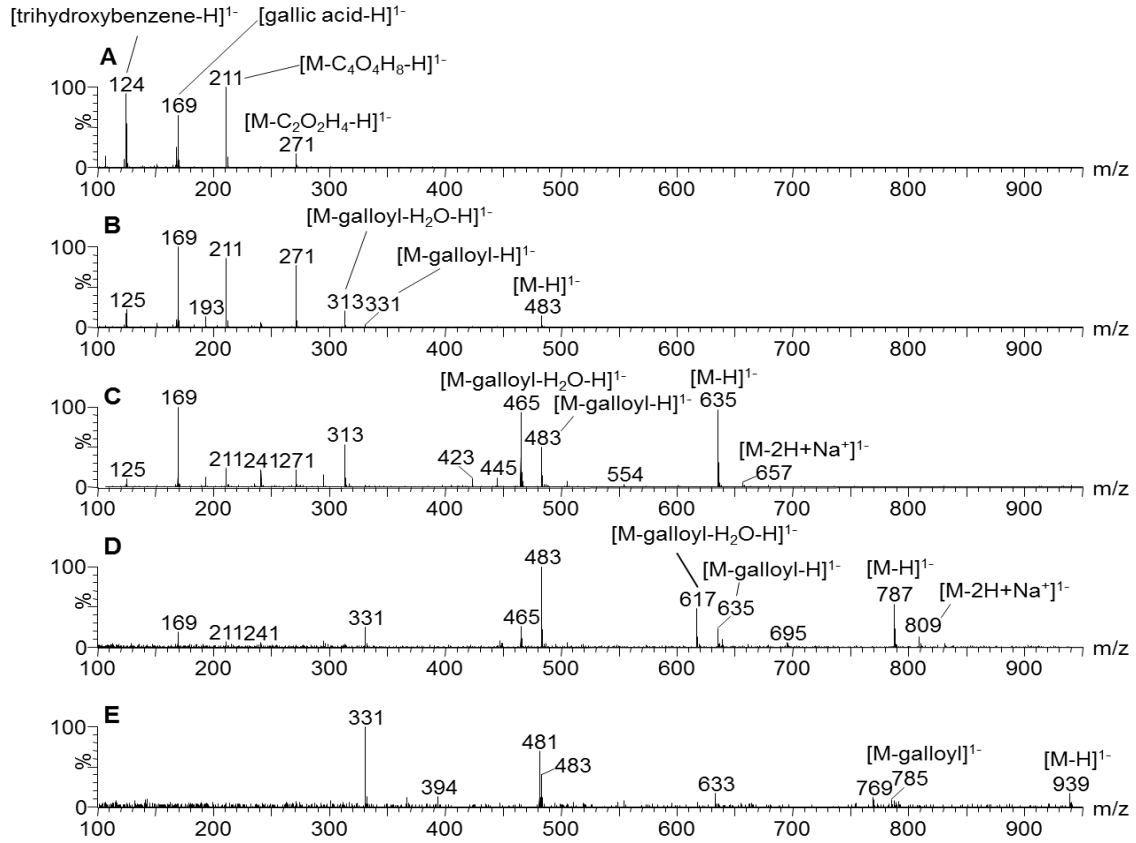


Figure S26: Representative MS^E spectra obtained for monogalloyl glucose (**A**, **2**), digalloyl glucose (**B**, **15**), trigalloyl glucose (**C**, **30**), tetragalloyl glucose (**D**, **33**) under HILIC conditions, and pentagalloyl glucose (**E**, **37**) obtained from RP-LC.

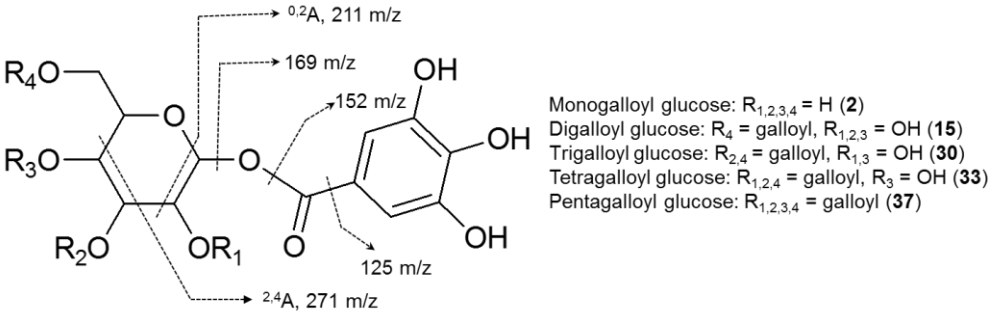


Figure S27: Molecular structures and fragmentation pathways for galloyl glucose species detected in chestnut. $^{0,2}A^-$ and $^{2,4}A^-$ ion masses are indicated for monogalloyl glucose (**2**).

Table S3: Comparison of CCS values of selected hydrolysable tannins obtained on a TWIMS ($^{TW}CCS_{N2}$) instrument in the current work with values determined on a DTIMS ($^{DT}CCS_{N2}$) instrument.

Compound (no.)	Ion Species	m/z	$^{TW}CCS_{N2}$ (Å^2)	$^{DT}CCS_{N2}$ (Å^2)	% Error
HHDP-Glucose (**3**)	$[M-H]^{1-}$	481.06	197.5	198.0	0.26
Pedunculagin - CO_2 (**4**)	$[M-H]^{1-}$	739.08	251.5	249.7	-0.73
Dihydrated Vescalagin (**5**)	$[M-H]^{1-}$	969.08	283.8	281.1	-0.95
Castalagin-castalin/vescalin dimer (**7**)	$[M-2H]^{2-}$	773.05	393.1	396.1	0.77
Galloyl-HHDP-glucose (**8**)	$[M-H]^{1-}$	633.07	221.6	221.7	0.02
Roburin A/D (**10**)	$[M-2H]^{2-}$	924.06	414.1	413.8	-0.07
			462.1	454.5	-1.67
Roburin B/C (**11**)	$[M-2H]^{2-}$	990.07	438.1	433.6	-1.04
C-galloylated vescalagin dimer (**12**)	$[M-2H]^{2-}$	1000.06	442.1	438.7	-0.77
			480.0	473.8	-1.29
Vescavaloninic acid (**16**)	$[M-H]^{1-}$	1101.07	318.8	313.7	-1.61
Vescalagin (**17**)	$[M-H]^{1-}$	933.07	290.1	287.8	-0.80
Vescalagin dimeric cluster	$[2M-2H]^{2-}$	933.06	392.2	388.7	0.90
O-galloylated castalagin/vescalagin (**19**)	$[M-H]^{1-}$	1085.07	314.8	310.2	-1.48
	$[M-2H]^{2-}$	542.03	338.3	338.4	0.02
1-O-trihydroxybenzene castalagin (**26**)	$[M-H]^{1-}$	1041.08	307.7	303.1	-1.50
	$[M-2H]^{2-}$	520.04	333.2	334.9	0.49
Tetragalloyl glucose (**33**)	$[M-H]^{1-}$	787.09	263.6	260.8	-1.04
	$[M-2H]^{2-}$	393.04	297.5	297.1	-0.12
Ellagic acid-pentose (**34**)	$[M-H]^{1-}$	433.04	191.0	193.0	1.05
Ellagic acid (**36**)	$[M-H]^{1-}$	300.99	147.2	154.8	5.16
Trimethoxy-ellagic acid (**38**)	$[M-H]^{1-}$	343.04	169.6	174.4	2.81

Declaration with signatures in possession of candidate and supervisor.

Declaration by the candidate:

With regard to Chapter 4, the nature and scope of my contribution were as follows:

Nature of contribution	**Extent of contribution (%)**
Performed the experiments, data analysis, co-wrote paper	70

The following co-authors have contributed to Chapter 4:

Name	**E-mail address**	**Nature of contribution**	**Extent of contribution (%)**
Harald Pasch	hpasch@sun.ac.za	Editorial input	5
André de Villiers	ajdevill@sun.ac.za	Data analysis, editorial input	25

Signature of candidate:

Date: 25\1\2018

Declaration by co-authors:

The undersigned confirm that:

1. The declaration above accurately reflects the nature and extent of the contributions of the candidate and the co-authors to Chapter 4,
2. No other authors contributed to Chapter 4 besides those specified above, and
3. Potential conflicts of interest have been revealed to all interested parties and that the necessary changes have been made to use the material in Chapter 6 of this dissertation.

Signature	**Institutional affiliation**	**Date**
	Stellenbosch University	25/1/2018
	Stellenbosch University	25/1/2018

Chapter 4

Comprehensive analysis of hydrolysable tannins by reversed phase and hydrophilic interaction chromatography coupled to ion mobility and high-resolution mass spectrometry, Part 2: Tara tannins

Abstract

In this, the second of a two-part study, reversed-phase LC (RP-LC) and hydrophilic interaction chromatography (HILIC) methods hyphenated to diode array detection and ion mobility (IM) - high resolution mass spectrometry (HR-MS) were used for the analysis of gallotannins in a commercial tara extract. UV spectra combined with low and high collision energy mass spectral data and known RP-LC elution orders allowed the identification of 43 isomeric gallotannins. The synergy between IM and UV data was found to provide a simple means to determine the number of depsidic bonds and thus to distinguish between positional isomers. IM also facilitated the assignment of individual isomeric species between HILIC and RP-LC separations. For the gallotannin species present in tara, RP-LC provided superior resolution and specificity compared to HILIC. The results reported in this paper confirm the utility of IM in combination with optimised complementary chromatographic separations and HR-MS for the detailed qualitative analysis of hydrolysable tannins in complex mixtures of these compounds.

4.1 Introduction

Tara polyphenols comprise specific gallotannins [1] composed of a quinic acid esterified to gallic acid moieties (galloyl-quinic acids (GQAs), **Figure 4.1**), which may also further form aryl ester(s) (depsides) with one or more additional gallic acid residues. A large number of GQA regioisomers which differ in the position(s) and nature of the respective acyl moieties exist in natural products. These compounds are structurally related to chlorogenic acids (caffeoyl-quinic acids (CQAs)), where the quinic acid core is esterified to caffeic acid units [1]. An important distinguishing feature of GQAs is the depsidic bond formed between gallic acid units in tara gallotannins [1, 2]. GQAs containing depsidic bonds show a characteristic shoulder at 300 nm in their UV spectra, which increases in intensity with the number of depsidic bonds [3].

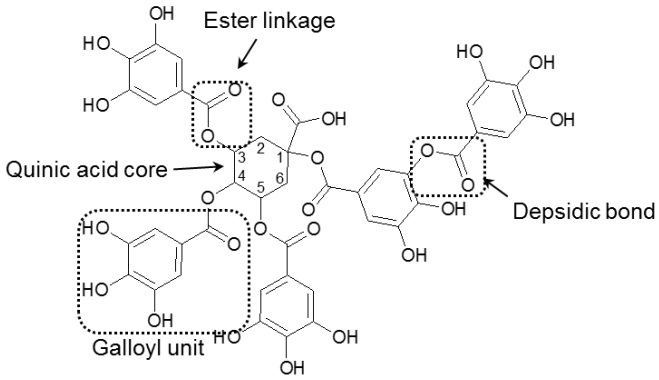


Figure 4.1: Representative structure of a gallotannin found in tara. For structures of the compounds identified in the current work, the reader is referred to **Figure S1**.

Tara tannins are mainly used for the tanning of animal hides used in car interiors, due to their properties conferring grain cracking resistance [4]. Other potential uses for tara tannins include their application as resins [5, 6], wood adhesives [7], as protection against steel oxidation [8], as rust removers [9] and as natural antibacterials [10].

Due to the large number of regioisomers comprising tara tannins, their analytical characterisation is challenging. Solvent extraction is commonly performed using water and methanol for low molecular weight (MW) gallotannins, and acetone, ethyl acetate

or ethanol for high MW gallotannins [11–13]. For the purification of gallotannins, fractionation of solvent extracts by size exclusion chromatography (SEC) on Sephadex LH-20 followed by semi-preparative C_{18} reversed phase (RP) isolation is typically used [11, 14, 15]. Normal phase liquid chromatography (NP-LC) fractionation followed by preparative isolation on a C18 column has also been used for the eventual NMR characterisation of gallotannins [16]. In terms of analytical separation, NP-LC has been used in the past [17, 18], but reversed phase liquid chromatography (RP-LC) is nowadays the preferred mode of HPLC for hydrolysable tannins in general [12].

Gallotannins show a typical UV absorbance band at 280 nm, with a bathochromic shift of 10 -12 nm depending on the degree of galloylation [12]. For mass spectrometric identification, negative electrospray ionisation is generally employed [12], with typical fragmentation involving successive neutral losses of galloyl groups (-152 amu), and both quinic acid (m/z 191) and gallic acid moieties (m/z 169) detected as fragment ions [19]. Clifford and co-workers [20, 21] reported a fragmentation scheme to discriminate between chlorogenic acid regioisomers, which was subsequently also applied to the identification of GQAs in tara [19]. The same relative RP-LC elution order was noted for both mono and di-acylquinic acids, regardless of the chemical nature of the acyl unit [19, 20, 22, 23].

Due to its ability to produce mainly singly charged ions, matrix-assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) has been used to determine the molecular weight distribution of protein-(tara)tannin complexes [24]. This work showed that in tara, quinic acid can be substituted with up to eight galloyl units. Pizzi *et al*. [25] hypothesised that tara tannins could include structures comprising an ellagic acid core esterified to gallic acid with an additional CO_2 unit, although the limited resolution of the TOF instrument used made distinction between these species and GQAs difficult.

More recently, ion mobility-MS (IM-MS) has also found application in the analysis of ellagitannins [26] and chlorogenic acids [27–29]. Ion mobility proved useful for the separation of chlorogenic acid regioisomers as well as prototropic isomers differing their deprotonation sites [28]. Zheng *et al*. [29] recently reported the use of drift tube ion mobility and a newly developed ultra high resolution IM system (SLIM, structures

for lossless ion manipulation) to successfully differentiate between *cis/trans* dicaffeoylquinic acid isomers.

In the present two-part study, we report an analytical approach based on hydrophilic interaction chromatography (HILIC) and RP-LC chromatographic methods hyphenated to diode array detection and travelling wave ion mobility (TWIMS)-HR-MS for the detailed analysis hydrolysable tannins. The application of this approach to ellagitannins in chestnut is reported in chapter 3. In the present work we report our findings for the analysis of tara gallotannins. The emphasis is on the use of these methods to distinguish between regioisomeric GQA species.

4.2 Experimental

4.2.1 Materials and reagents

The tara tannin sample, an acetone-water extract of Peruvian tara (*Caesalpinia spinosa*) pods, was obtained from Silvateam (San Michele Mondovì, Italy). HPLC grade acetonitrile, methanol (MeOH) and formic acid were purchased from Sigma-Aldrich (Johannesburg, South Africa) and deionised water was obtained using a Milli-Q water purification system (Millipore, Milford, MA). Poly-DL-alanine used for IM calibration was purchased from Sigma-Aldrich.

4.2.2 Instrumentation and chromatographic conditions

Analyses were performed on an Acquity UPLC system hyphenated to a photodiode array (PDA) detector (500 nL flow cell, 10 mm path length) and a Synapt G2 quadrupole time-of-flight (Q-TOF) mass spectrometer equipped with an ESI source operated in negative ionisation mode (Waters, Milford, MA, USA). UV and MS detection were performed as specified in chapter 3. Briefly, the MS scan range was 100-2000 amu at a scan time of 0.2 seconds. IM measurements were performed using N_2 as drift gas at 90 mL/min and mobility T-Wave velocity and wave height of 448 m/s and 37.1 V, respectively.

Collisional cross sections (CCSs) were calculated according to [30] using poly-DL-alanine as calibrant [31] as outlined in chapter 3.The accuracy of measured $^{TW}CCS_{N2}$

values were confirmed for selected compounds by comparison with $^{DT}CCS_{N2}$ values determined on a drift tube IM instrument (Agilent 6560 Ion Mobility Q-TOF) (**Table S3** in chapter 3).

For both RP-LC and HILIC separations acidified water (0.1% formic acid) was used as mobile phase A and acetonitrile as mobile phase B. The tannin sample was prepared by dissolving 12 mg/mL in MeOH. 2 µL was injected in full loop mode using acetonitrile and MeOH/H_2O (50/50, v/v) as weak and strong needle washes, respectively.

4.2.3 HILIC analyses

HILIC separations were performed on an XBridge Amide (150 × 4.6 mm i.d., 2.5 µm) column (Waters) by using a linear gradient of 10-12.5% A (0-20 min) at a flow rate of 1 mL/min. The column effluent was split 1:3 between the PDA and MS detectors.

4.2.4 RP-LC analyses

Separations were performed on a Kinetex C_{18} (100 × 2.1 mm i.d., 1.7 µm) superficially porous column (Phenomenex, Torrance, USA) using the following linear gradient: 2-20% B (0.0–40 min), 20% B (40–50 min). The flow rate was 0.4 mL/min, with the total flow directed to the MS source.

4.2.5 Data processing

Data acquisition and processing were performed using MassLynx (v. 4.1) and Driftscope (v. 2.1) software (Waters). Reported mass spectra were filtered as a function of arrival time for both low and high collision energy data using Driftscope.

4.3 Result and discussion

4.3.1 Chromatographic separation

Base peak ion (BPI) chromatograms obtained for the RP-LC- and HILIC-IM-MS analyses of the tara tannin extract are presented in **Figure 4.2**, where numbers specify the number of galloyl units and superscript letters distinguish isomeric species**.** For the gallotannins present in tara, unlike the hydrolysable tannins in chestnut (chapter 3), HILIC provided relatively poor chromatographic performance. Furthermore, for the

GQA derivatives, the general HILIC retention order (i.e. increasing retention with increasing degree of galloylation) was similar to that obtained by RP-LC, since a higher number of gallic acid moieties increases both the hydrophobicity (due to the aromatic groups) and polarity (due to hydroxyl groups). However, for regioisomers of the same degree of galloylation, the relative elution order often differs between the separation modes (compare the elution orders of compounds with the same number according to superscripts).

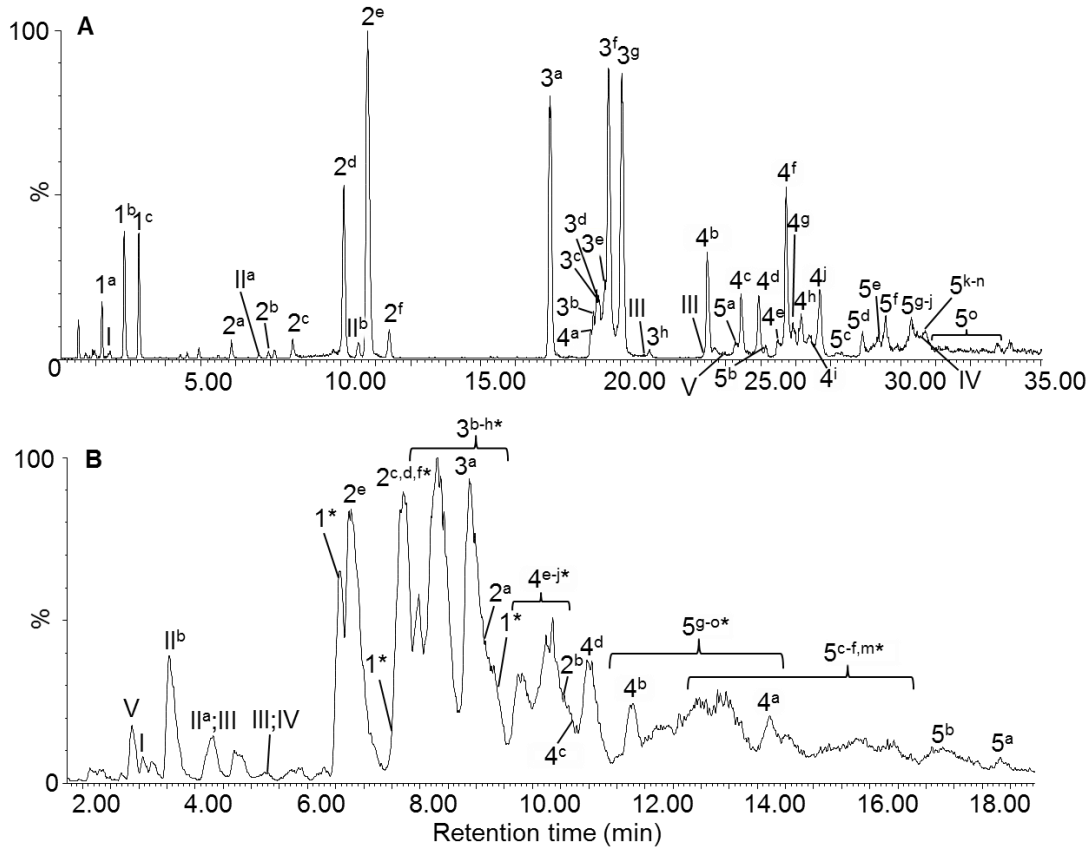


Figure 4.2: Base peak ion chromatograms obtained for the (**A**) RP-LC- and (**B**) HILIC-IM-Q-TOF-MS analyses of tara hydrolysable tannins. Peak labels correspond to **Table 4.1** and **Figure S1**; numbers specify the number of galloyl units for each compound and superscript letters distinguish between isomeric species. *Indicates isomers which could not be matched between RP-LC and HILIC. For experimental conditions, refer to **Section 4.2.2**.

4.3.2 Identification of tara hydrolysable tannins

For the tentative identification of compounds **1**[a-c], **2**[a-f] ,**3**[a] ,**4**[a] and **5**[b] in tara (**Table 4.1** and **Figure S1**), RP-LC elution orders, UV spectral information and high resolution high energy MS data were compared to previous literature reports [19, 20, 22, 23, 32–34]. UV spectra are useful in the differentiation of GQA structural isomers of the same degree of galloylation, but differing in the number of depsidic bonds, based on the relative absorbance at 300 nm due to the depsidic bond [3]. Concerning RP-LC retention, previous studies have shown that the structurally related monoacyl chlorogenic acids (CQAs) elute in the sequence 3-, 5- and 4-O-caffeoyl-quinic acid [32, 33]. The same elution order has been reported for the positional isomers of monoacyl GQAs in RP-LC [19, 20]. Also for diacylated CQAs and GQAs, an identical RP-LC elution order is observed in the sequence 3,4-, 3,5- and 4,5-diacyl-quinic acids, irrespective of the nature of the acyl group [19, 20, 22, 23].

The substitution position of the caffeoyl moiety on the quinic acid core in mono-acylated CQAs can also be identified by MS^n [20, 34]. High collision energy fragmentation spectra show a distinctive dehydrated quinic acid ion at fragment 173 m/z, ([quinic acid-H_2O-H]$^-$) when the acyl unit is located on the 4-position (4-CQA), whereas 3-CQA is distinguished from the 5-CQA positional isomer by a comparatively more intense hydroxycinnamic acid (179 m/z, [caffeic acid-H]$^{1-}$) ion relative to the 191 m/z ([quinic acid-H]$^{1-}$) base peak ion. On this basis, Clifford *et al*. [20] derived a protocol to differentiate diacyl CQAs based on MS^n data. 3,5-di-CQA is easily identified as this compound does not show a dehydrated quinic acid ion at 173 m/z, whereas 3,4-di-CQA is differentiated from 3,5-di-CQA by the relative intensity of the caffeic acid ion in the MS^3 spectrum. The likelihood of intact hydroxycinnamic acid ions being formed during fragmentation decreases in the sequence 1 ≈ 5 > 3 > 4 for substituted CQAs [20, 21], with the facile loss of the C5 acyl unit rationalised [21] as being dependent on the site of the negative charge on the precursor ion. Significantly, this behavior of acylated quinic acid derivatives is independent of the nature of the acyl moiety [19], such that the same protocol could be used to identify GCA derivatives in the present work.

Based on the criteria outlined above, a total of 37 GCAs comprising mono- to pentagalloylated quinic acid positional and structural isomers were identified in tara (**Table 4.1** and **Figure S1**), as discussed in more detail in the following sections.

Three monogalloyl-quinic acid species were identified based on their elution order and MS^E spectra under RP-LC conditions [19]. The last eluting isomer at 2.80 min (**1^c**) was identified as 4-O-galloylquinic acid, based on the characterictic 173 m/z fragment ion in its MS^E spectrum and its relative hydrophobicity. The two remaining isomers eluting at 1.15 and 2.28 min were identified as 3-O-galloyl- (**1^a**) and 5-O-galloylquinic acid (**1^b**), respectively, taking into acount their elution order as well as the low intensity acyl fragment (169 m/z) for **1^b**. A clean MS^E spectrum could not be obtained for 3-O-galloylquinic acid under either RP-LC or HILIC conditions due to co-elution, and this compound was therefore solely identified based on RP-LC elution order.To assign the monogalloyl-quinic acids (**1^{a,b,c}**) in HILIC, the relative peak areas between the two separations were used, which revealed the expected reversed elution order.

Table 4.1: Galloyl-quinic acids and gallic acids tentatively identified in tara by RP-LC- and HILIC-IM-HR-MS.

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
3-O-galloyl quinic acid (1[a])#	$C_{14}H_{16}O_{10}$	1.15	9.48	co-elution	343.0674	2.62	2.21	170	co-elution*
5-O-galloyl quinic acid (1[b])#	$C_{14}H_{16}O_{10}$	2.28	7.62	co-elution	343.0670	1.46	2.21	170	343.0836, 191.0649, 169.0211
4-O-galloyl quinic acid (1[c])#	$C_{14}H_{16}O_{10}$	2.80	6.59	273	343.0668	0.87	2.21	170	343.0831, 191.0647, 173.0524, 169.0208, 125.0291
3,4-di-O-galloyl quinic acid (2[a])	$C_{21}H_{20}O_{14}$	6.10	6.77, 7.71, 9.01, 10.70#	co-elution	495.0784	1.82	3.40	209	495.0770†, 343.0667, 331.0652, 325.0563, 191.0550, 173.0443, 169.0139
3,5-di-O-galloyl quinic acid (2[b])	$C_{21}H_{20}O_{14}$	7.12		co-elution	495.0777	0.40	3.40	209	343.0679, 331.0694, 191.0540, 169.0130†, 125.0237
3-O-digalloyl quinic acid (2[c])	$C_{21}H_{20}O_{14}$	8.29		281[sh]*	495.0775	0.00	3.10 3.60	199 215	495.0770, 365.0497, 343.0667, 191.0550†, 169.0143, 125.0219
5-O-digalloyl quinic acid (2[d])	$C_{21}H_{20}O_{14}$	10.12		272[sh]	495.0794	3.84	3.10 3.60	199 215	495.0782, 365.0498, 343.0672, 325.0568, 191.0550†, 169.0137
4,5-di-O-galloyl quinic acid (2[e])	$C_{21}H_{20}O_{14}$	10.97		276	495.0792	3.43	3.15	201	495.0793†, 343.0669, 325.0562, 191.0550, 173.0449, 169.0133, 125.0233
4-O-digalloyl quinic acid (2[f])	$C_{21}H_{20}O_{14}$	11.74		270[sh]	495.0776	0.20	3.30 3.70	206 218	495.0761, 365.0478, 343.0654†, 191.0565, 173.0430, 169.0131, 125.0225
3,4,5-tri-O-galloyl quinic acid (3[a])	$C_{28}H_{24}O_{18}$	17.48	8.90	275[sh]	647.0887	0.46	4.00	226	647.0939†, 495.0818, 477.0698, 343.0674, 325.0574, 169.0139

Table 4.1 (continued).

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
Trigalloyl quinic acid (one depsidic bond) (3^{b-h})	$C_{28}H_{24}O_{18}$	19.03, 19.14, 19.22, 19.44, 19.57, 20.05, 21.05	7.98, 8.30, 9.31#	276sh	647.0903	2.94	4.35	236	647.0891, 495.0787†, 343.0666, 325.0561, 191.0550, 169.0131
				273sh	647.0904	3.09	4.35	236	
				275sh	647.0908	3.71	4.35	236	
				273sh	647.0890	0.93	4.35	236	
1,3,4,5-tetra-O-galloylquinic acid (4^a)	$C_{35}H_{28}O_{22}$	18.94	14.39	275	799.0992	0.25	4.76	246	799.1014†, 647.0888, 629.0783, 601.0839, 495.0776, 477.0672, 191.0532, 169.0118
Tetragalloyl quinic acid with one depsidic bond (4^b)	$C_{35}H_{28}O_{22}$	23.09	11.79	274sh	799.1028	4.25	5.04	253	799.0979, 647.0874†, 495.0767, 169.0125
Tetragalloyl quinic acid with one depsidic bond (4^c)		24.31	10.57		799.1012	2.25	5.11	255	
Tetragalloyl quinic acid with one depsidic bond (4^d)		24.93	11.06		799.0999	0.63	5.24	258	
Tetragalloyl quinic acid (two depsidic bonds) (4^{e-j})	$C_{35}H_{28}O_{23}$	25.62, 25.91, 26.15, 26.45, 26.78, 27.11	9.46, 9.81, 10.36	272sh	799.0996	0.25	5.66	268	799.1007, 647.0910†, 495.0794, 343.0674,173.0435, 169.0128
Pentagalloyl quinic acid (two depsidic bonds) (5^a)	$C_{42}H_{32}O_{26}$	24.28	18.31	275sh	951.1125	2.21	6.28	282	951.1108, 799.0971, 647.0881†, 495.0754, 477.0647, 325.0558, 169.0124

Table 4.1 (continued).

Compound (no)	Chemical formula	t_R RP-LC (min)	t_R HILIC (min)	UV (nm)	Exp. m/z $[M-H]^{-1}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Fragment ions $[M-H]^{1-}$
1-O-digalloyl-3,4,5-tri-O-galloylquinic acid (5[b])	$C_{42}H_{32}O_{26}$	25.19	17.06	275[sh]	951.1116	1.26	5.80	282	951.1115, 799.1005, 647.0894, 629.0779, 601.0848, 495.0775, 477.0696, 325.0533, 169.0138†
Pentagalloyl quinic acid (two depsidic bonds) (5[c-f, m])	$C_{42}H_{32}O_{26}$	27.70, 28.65, 29.20, 29.46, 30.80	12.98, 13.42, 14.28, 18.31	274[sh]	951.1107	0.32	6.28	282	951.1152, 799.1005†, 647.0899, 495.0758, 323.0408, 173.0442, 169.0131
Pentagalloyl quinic acid (three depsidic bonds) (5[g-o])	$C_{42}H_{32}O_{26}$	30.40, 30.80, 30.90	11.72-12.28	273[sh]	951.1104	0.00	7.04	299	951.1085, 799.0965†, 647.0848, 495.0755, 323.0396, 173.0442, 169.0117
Gallic acid (I)	$C_7H_6O_5$	1.72	3.06	271	169.0137	2.96	0.83	107	125.0232
***p*-Digallic acid (II[a])[#]**	$C_{14}H_{10}O_9$	7.10	4.32	270[sh]	321.0247	1.56	2.00	162	125.0290, 169.0205
***m*-Digallic acid (II[b])[#]**		10.62	3.53						
Trigallic acid (III)[#]	$C_{21}H_{14}O_{13}$	20.78	5.15	269[sh]	473.0360	0.42	3.17	202	321.0377, 169.0211, 125.0306
		22.96	4.30				3.17	202	
		n.d.	4.60				2.90	195	
Tetragallic acid (IV)	$C_{28}H_{18}O_{17}$	31.14	5.27	co-elution	625.0466	0.80	4.21	232	co-elution
Ellagic acid (V)	$C_{14}H_6O_8$	23.68	2.90	253, 364	300.9987	-1.00	1.66	147	283.9947, 257.0065, 245.0087, 229.0147, 217.0122, 201.0180, 173.0251, 145.0271

*t_d: arrival time; co-elution: no clear fragments obtained due to co-elution; sh: shoulder.

[#]Individual isomeric species were assigned according to their repective peak areas.

†Base peak ion in the MS^E spectrum.

Assigning mono-galloylquinic acid species between RP-LC and HILIC based on MS^E data revealed that the protocol [20] originally developed by Clifford and co-workers for the assignment of CQAs under RP conditions is not suitable as a basis to identify the position of the galloyl unit under HILIC conditions. Comparison of MS^E spectra for the two major mono-galloylquinic acid isomers obtained under RP-LC (**Figures S2**) and HILIC (**Figure S3**) conditions shows that in the latter case, similar fragment ion intensities are observed, while the 173 m/z fragment ion, which is diagnostic for all 4-substituted quinic acid derivatives under RP-LC conditions, is also observed for both species. Clifford *et al*. [21] proposed that the fragmentation pathways followed by regioisomeric caffeoyl-quinic acids depend on the position of the negative charge under RP-LC conditions. If this is correct, the apparent loss in mass spectrometric specificity for the mono-galloylquinic acid species in HILIC may be a consequence of the negative charge being located on the same position for all three regiosiomers, possibly due to a lower polarity [35] or higher effective pH [36, 37] of the organic-rich HILIC mobile phase.

A total of six digalloyl-quinic acid derivatives were detected and identified based on accurate mass information. Isomers eluting at 8.29, 10.12 and 11.74 min in RP-LC are distinguished by a shoulder at 300 nm in their UV spectra, which points to these compounds containing a depsidic bond and therefore only one site of galloylation. These three regioisomers were characterised by multiple arrival times at 3.10-3.30 and 3.60-3.70 ms (refer to **Figure 4.6** and discussion below). The last eluting isomer (**2^f**, 11.74 min in RP-LC) shows the characteristic dehydrated quinic acid fragment ion at m/z 173, identifying this compound as 4-O-digalloyl quinic acid (**Figure S4-C**). The other two isomeric species eluting at 8.29 and 10.12 min were therefore identified as 3-O- (**2^c**, **Figure S4-A**) and 5-O-digalloyl quinic acid (**2^d**, **Figure S4-B**), respectively, taking into account the elution order of their respective monogalloyl derivatives. Two of the three mono-depsidic digalloyl-quinic acid isomers co-eluted in HILIC, and fragmentation spectra revealed no differences for the peaks at retention times of 7.71 and 9.01 min (data not shown), so that these isomers could not be distinguished in this separation mode.

For the digalloyl-quinic acid isomers not containing a depsidic bond, the isomer eluting at 7.12 minutes in RP-LC was identified as 3,5-di-O-galloyl quinic acid (**2^b**) based on

the absence of the dehydrated quinic acid fragment ion (m/z 173) in its MS^E spectrum (**Figure S5-B**). The two remaining isomers could not be differentiated based on the available MS data. However, taking into account the relative elution order for these isomers reported on a phenyl-hexyl RP column [20], the isomers eluting at 6.10 and 10.97 min in RP-LC were tentatively identified as 3,4-di-O-galloyl- (**2ª**, **Figure S5-A**) and 4,5-di-O-galloyl-quinic acid (**2ᵉ**, **Figure S5-C**), respectively.

Eight trigalloyl-quinic acid derivatives were identified based on HR-MS data. The first isomer to elute under RP-LC conditions showed no shoulder at 300 nm in its UV spectrum, identifying this component as 3,4,5-tri-O-galloyl quinic acid (**3ª, Figure S6-1**). The later eluting isomers showed an increase in arrival time in IM from 4.00 to 4.35 ms, and a shoulder at 300 nm in their UV spectra (**Figure S6-2**). Comparing the HILIC×IM (**Figure S6-A**) and RP-LC×IM (**Figure S6-B**) contour plots allowed for straightforward identification of 3,4,5-tri-O-galloyl-quinic acid in HILIC due to its unique arrival time at 4.00 ms. The MS^E spectra (**Figure S7**) of this isomer in both HILIC and RP-LC are unique due to the presence of an extra fragment ion at 477 m/z ([M-galloyl-H_2O-H]$^{1-}$), which is also evident in the high energy spectrum of 1,3,4,5-tetra-O-galloyl-quinic acid (**4ª**, see further) [19]. The remaining trigalloyl species (**3ᵇ⁻ʰ**) all produced the same fragmentation spectra in RP-LC (**Figure S8-A**). Since seven of these isomers were resolved by RP-LC compared to only three in HILIC, and identical arrival times were measured for these, assignment of individual species between HILIC and RP-LC was not possible.

The MS^E spectra obtained for the GQAs in tara generally showed more extensive fragmentation under HILIC compared to RP-LC conditions - as indeed observed for chestnut hydrolysable tannins (chapter 3) - which may prove beneficial in some cases. For example (different from the case for mono-galloylquinic acid species) the diagnostic intensities of ions 169, 173 and 191 m/z for the tri-galloylquinic acid isomer eluting at 9.30 min in HILIC allowed for identification of the galloylation positions as 3 and 5 due to the absence of the 173 m/z peak, although the position of the depsidic bond could not be determined (**Figure S8-C**). This observation suggests that when comparable MS data between RP-LC and HILIC are required, such as for library searching for example, higher collision energies should be used in RP-LC.

A total of ten tetragalloyl-quinic acid derivatives were detected in the tara extract. The isomer eluting first in RP-LC was easy to assign as 1,3,4,5-tetra-O-galloylquinic acid (**4ª**) by the absence of a shoulder at 300 nm in its UV spectrum and the fact that this compound showed the shortest arrival time of 4.76 ms (refer to **Figure 4.3** below). This is further supported by the unique fragmentation behaviour of this isomer, which exhibits an ion at 601 m/z ([M-galloyl-CO_2H-H]$^{1-}$, 0.5 ppm) under MS^E conditions (**Figure S9)** [19].

IM provided valuable complementary information to RP-LC and HILIC separations regarding the number of depsidic bonds, which, together with HR-MS^E data, greatly facilitated the tentative identification of positional isomers. As an example, the HILIC- and RP-LC × ion mobility extracted ion contour plots for tetragalloyl-quinic acid species (m/z 799, [M-H]$^-$) are shown in **Figure 4.3**. 1,3,4,5- tetra-O-galloyl quinic acid (**4ª**, **Figure 4.3-1**), identified as discussed above, showed the shortest arrival time of all tetragalloyl-quinic acid derivatives at 4.76 ms. Two additional arrival time clusters where recorded at 5.04-5.24 and 5.66 ms, respectively (**Figure 4.3-C**), both coinciding with an increasingly evident shoulder at 300 nm in the UV spectra of the respective peaks (**Figure 4.3-2,3**). (A similar increase in arrival time and relative absorbance at 300 nm as a function of number of depsidic bonds was observed for tri- and penta-galloyl-quinic acid species, **Figures S6** and **4.5**, respectively). This indicates, as expected, that the effective $^{TW}CCS_{N2}$ values of GQA derivatives increase with the number of depsidic bonds (**Table 4.1**). The complementary information provided by UV and IM data, therefore, provides a simple means of distinguishing between positional isomers containing different numbers of depsidic bonds. Furthermore, the characteristic arrival times for particular positional isomers facilitated assignment of individual isomeric species of di-, tri-, tetra- and penta-galloylquinic acids between HILIC and RP-LC (**Figures 4.6, S6, 4.3** and **4.5,** respectively), which generally showed an inverse elution order. For example, minor differences in arrival time were measured for regioisomeric tetragalloyl-quinic acid species (**Figure 4.3**) containing one depsidic bond (**4ᵇ**, **4ᶜ** and **4ᵈ**) indicate that **4ᶜ** and **4ᵈ** retain the same elution order in HILIC and RP-LC. Although no further information regarding the substitution patterns could be obtained, IM offered partial differentiation between these isomeric species, which was not possible based on MS data alone (**Figure S10**).

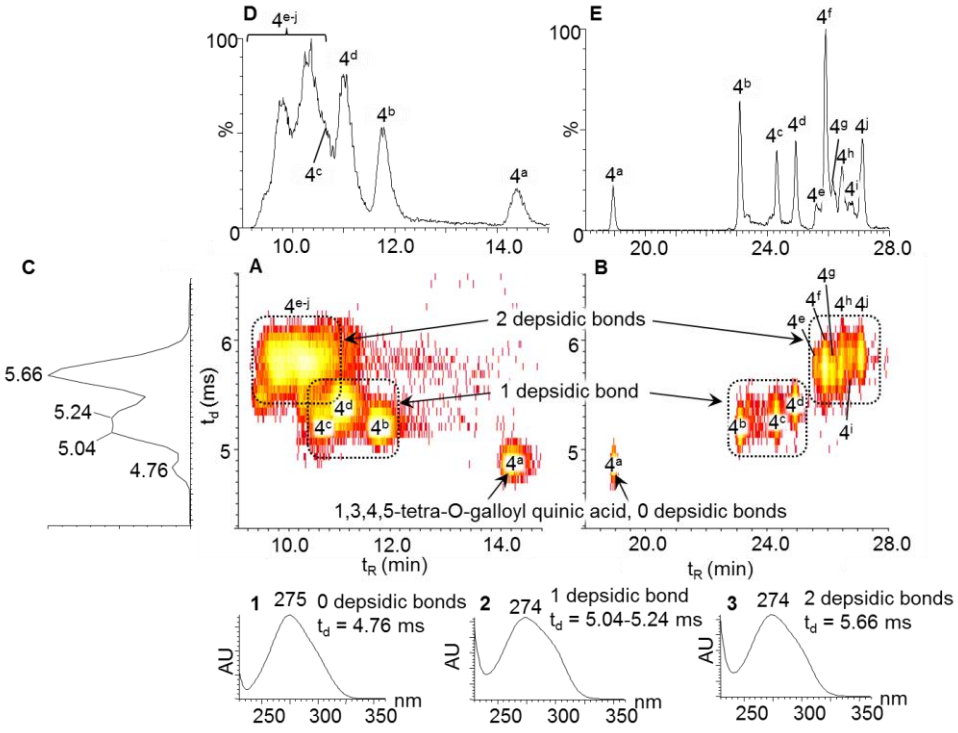


Figure 4.3: Extracted ion HILIC×ion mobility (**A**) and RP-LC×ion mobility (**B**) contour plots for tetragalloyl-quinic acid derivatives (m/z 799) in tara. (**C**) shows the corresponding extracted ion arrival time plot (identical for HILIC and RP-LC), (**D**) and (**E**) the corresponding extracted ion chromatograms for HILIC and RP-LC and inserts **1**-**3** representative UV spectra of each of the positional isomers containing 1, 2 or 3 depsidic bonds, respectively. Compound numbers correspond to **Table 4.1**.

Based on the criteria outlined above, tetragalloyl-quinic acid derivatives **4^{b-d}** could be assigned as containing a single depsidic bond, while compounds **4^{e-j}** contained two depsidic bonds. Unlike **4^{b-d}**, however, the isomers containing two depsidic bonds showed nearly identical arrival times (5.66 ms) and fragmentation behaviour, and, especially in HILIC, extensive co-elution. As a consequence, assignment of individual isomers was not possible for these isomers.

Fifteen pentagalloyl-quinic acid derivatives were identified by HR-MS (m/z 951, $C_{42}H_{31}O_{26}$). The pentagalloyl-quinic acid isomers eluting at 24.28 (**5^{a}**) and 25.19 (**5^{b}**) min in RP-LC and 17.06 (**5^{b}**) and 18.31 (**5^{a}**) min in HILIC produce a characteristic ion at m/z 601 ([M-2galloyl-CO_2H-H]$^{1-}$, 0.5 ppm) under MS^{E} conditions (**Figure 4.4**),

indicating that an O-digalloyl group containing one depsidic bond is attached to quinic acid at position 1 for both isomers [19]. The formation of fragment ions 647 m/z ([M-2galloyl-H]$^{1-}$), 601 m/z ([M-2galloyl-CO_2H-H]$^{1-}$), and 629 m/z ([M-2galloyl-H_2O-H]$^{1-}$) can be explained by an acyl transfer mechanism (**Figure S11**, adapted from scheme B in [21]). However, no further information regarding the nature of the remainder of acyl groups could be obtained by either MS or UV data. The key to distinguishing between these isomers is the longer arrival time of **5^{a}** (6.28 vs. 5.80 ms for **5^{b}**), which indicates that this compound contains an additional depsidic bond, although its position could not be ascertained (**Figure 4.5**). Based on this information, **5^{b}** was tentatively identified as 1-O-digalloyl-3,4,5-tri-O-galloylquinic acid, while **5^{a}** contains a second depsidic chain at position 3,4 or 5 on the quinic acid core.

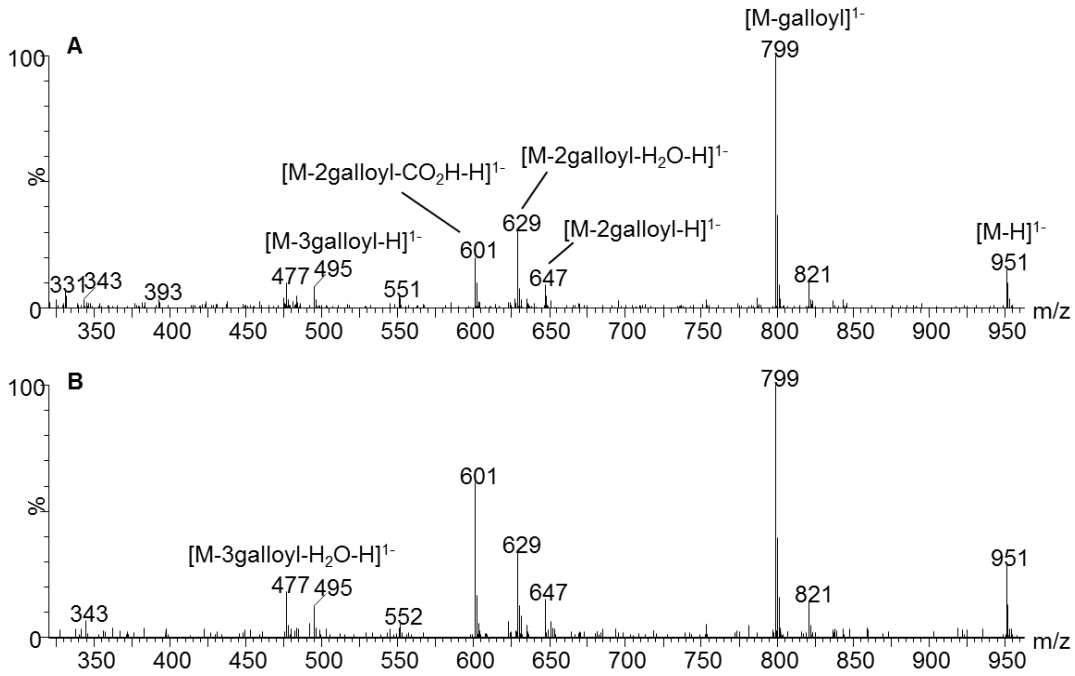


Figure 4.4: MS^{E} spectra obtained for pentagalloyl-quinic acid isomers eluting at 17.06 (**A**, **5^{b}**) and 18.31 (**B**, **5^{a}**) minutes under HILIC conditions.

The fragmentation spectra for the remaining pentagalloylquinic acid species at arrival times of 6.28 and 7.04 ms, representing isomers containing 2 and 3 depsidic bonds, respectively, were very similar (**Figure S12**). The only noticeable difference is the higher abundance of the fragment ion at 495 m/z [M-3galloyl-H]$^{-}$ in the MS^{E} spectra of

the species containing 3 depsidic bonds (t_d = 7.04 ms), indicating that cleavage of depsidic chains occurs more readily than single O-galloyl groups.

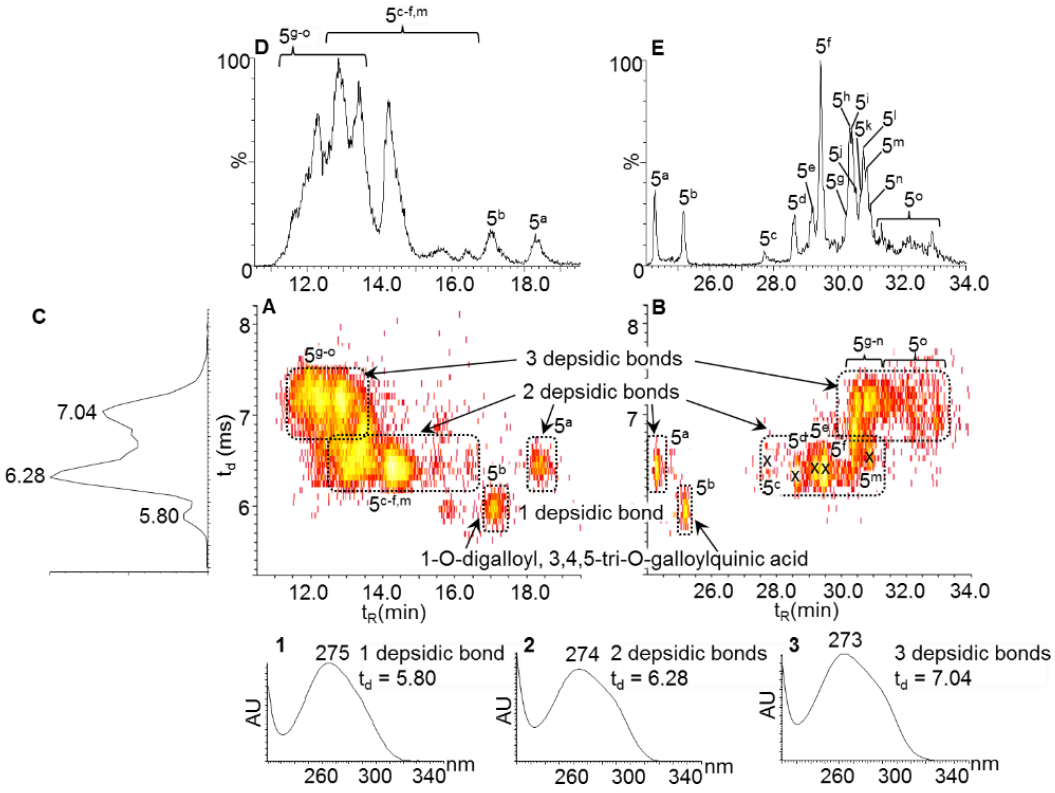


Figure 4.5: Extracted ion HILIC×ion mobility (**A**) and RP-LC×ion mobility (**B**) contour plots for pentagalloyl-quinic acid derivatives in tara. (**C**) shows the corresponding extracted ion arrival time plot (identical for HILIC and RP-LC), (**D**) and (**E**) the corresponding extracted ion chromatograms for HILIC and RP-LC and inserts **1**-**3** representative UV spectra of positional isomers containing 1, 2 or 3 depsidic bonds, respectively. Compound numbers correspond to **Table 4.1**.

Multiple arrival times (3.10~3.30 ms and 3.60~3.70 ms) were observed for each of the digalloyl quinic acid isomers containing a single depsidic bond (**2^c**, **2^d** and **2^f**), as shown in **Figure 4.6**. Since the positions of the depsidic bonds for each of these compounds has been ascertained based on MS data, the two species resolved for each compound by IM do not correspond to positional isomers. Rather, the species resolved by IM likely correspond to either prototropic isomers differing in the position of their negative

charges [28, 35] or to structural isomers containing *m*- and *p*-depsidic bonds, respectively [17]. MS^E spectra for each of the IM-resolved species filtered according to arrival time are presented in **Figure S13**. These indicate no discernible differences between the species separated by IM, with the exception of the higher prevalence of sodium and potassium adducts for the 4- and especially the 5-O-digalloyl quinic acid species with the later arrival times (**Figure S13**, **B1** and **C1**).

Considering that the *m*- and *p*-isomers of digallic acid are quite well separated in both HILIC and RP-LC (**Figure 4.2**, **II**^{a,b}) and that IM did not provide differentiation of these species (see below), the fact that no chromatographic separation of the IM-resolved digalloyl quinic acid species is evident in either separation mode implies that these more likely correspond to prototropic isomers. However, this could not be conclusively proven based on the available data.

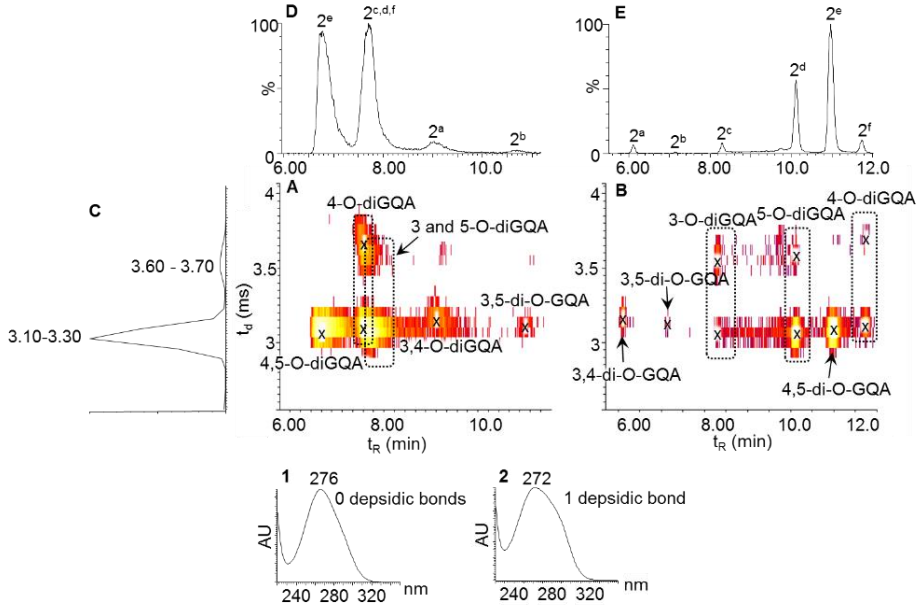


Figure 4.6: Extracted ion RP-LC×ion mobility (**A**) and HILIC×ion mobility (**B**) contour plots for digalloyl-quinic acid (diGQA) derivatives in tara. **C**) shows the corresponding extracted ion arrival time plot (identical for HILIC and RP-LC), (**D**) and (**E**) the corresponding extracted ion chromatograms for HILIC and RP-LC and inserts **1** and **2** representative UV spectra of positional isomers containing one and no depsidic bonds, respectively. Compound numbers correspond to **Table 4.1**.

In general, the results reported in the present work for galloyl quinic acids compare well with a previous study of structurally related chlorogenic acid regioisomers [29] in terms of relative IM behaviour. 4,5-di-O-caffeoyl quinic acid and the structurally related 4,5-di-O-galloyl quinic acid (**2^{e}**) in tara both showed shorter arrival times relative to the respective 3,4- and 3,5-di-O-acyl-quinic acid species, which made it possible to assign individual non-depsidic digalloyl-quinic acid species between HILIC and RP-LC (**Figure 4.6**). It was straightforward to assign the depsidic digalloyl-quinic acid isomers in HILIC compared to RP-LC due to the multiple arrival times observed for the proposed prototropic isomers, which subsequently indicated all three depsidic isomers (**2^{c}**, **2^{d}** and **2^{f}**) co-elute in HILIC (**Figure 4.6**).

For several of the O-di-, tetra- and penta-galloyl quinic acid derivatives (**2^{c,d}**, **4^{b,c,d}**, **5^{a,b}**) evidence of on-column conversion between species was observed in RP-LC, as reflected by a characteristic raised baseline (peak–plateau-peak) between species in the relevant extracted ion chromatograms (**Figure S14**). This phenomenon is well known for chiral compounds [38, 39], and has been observed previously, although not addressed in detail, for tara polyphenols [19]. Although not clearly evident in HILIC, this might be due to the relatively poor chromatographic performance of the technique. These inter-conversion reactions likely correspond to on-column conversions between the relevant positional isomers (e.g. between 3- and 5-O-digalloyl-quinic acid (**2^{c}** and **2^{d}**), and between tetragalloylquinic acid species with one depsidic bond (**4^{b,c,d}**). The pentagalloyl-quinic acid derivatives **5^{a}** and **5^{b}** with unique arrival times at 6.28 and 5.80 ms also show this characteristic raised baseline due to on-column conversion (**Figure S14-C**). If **5^{b}** (t_d = 5.80 ms) is indeed 1-O-digalloyl-3,4,5-tri-O-galloylquinic acid, which converts to **5^{a}** (t_d = 6.28 ms) containing two depsidic bonds. This provides evidence that galloyl units are not only capable of migrating to different positions on the quinic acid core, but also to phenolic groups on (presumably adjacent) galloyl units. In IM, the ‘interconverting’ species representing the raised baseline not surprisingly showed the same arrival time as the corresponding isomers (**Figure 4.6**). The observation of on-column conversion may therefore further complicate assignment of positional isomers, depending on the efficiency of the chromatographic separation.

Finally, several gallic acid derivatives were also detected in tara, including gallic acid (**I**), digallic acid (**II**), trigallic acid (**III**), and tetragallic acid (**IV**) (**Figure S1**). The

presence of these compounds is likely a consequence of the hydrolysis of GQAs. The di- and tri-gallic acid species showed the expected shoulder at 300 nm in their UV spectra due to the depsidic bond(s) (it was not possible to obtain a clean UV spectrum for tetragallic acid due to co-elution). These compounds were characterised by simple fragmentation patterns, with MS^E spectra showing losses of galloyl groups and both deprotonated (m/z 169) and decarboxylated (m/z 125) gallic acid ions.

Two chromatographic peaks were observed for both digallic acid (**II**) and trigallic acid (**III**) in both HILIC and RP-LC. Identical arrival times were measured for these species (**Table 4.1**). For digallic acid, the two peaks correspond to *m*- and *p*-depsidic isomers, with *p*-digallic acid eluting before *m*-digallic acid in RP-LC [40]. It has been shown that *m*-digallic acid is converted to *p*-digallic acid in methanol [17]. Since the tara sample was dissolved in methanol, such a conversion may be responsible for the detection of *p*-digallic acid in this work. Furthermore, oxidation of *m*-digallic acid has been shown to results in the formation of ellagic acid (**V**) [41]. This compound was also detected in the tara sample (**Table 4.1**), and was identified based on relative RP-LC and HILIC retention, MS^E and UV spectra and arrival time in accordance with chapter 3. For trigallic acid, three of the four possible isomers (*m-p*, *p-m*, *m-m* and *p-p*) were observed in HILIC, and only two in RP-LC. These two peaks could be assigned between the separation modes based on their relative peak areas, since both showed identical arrival times. The third trigallic acid species, detected only in HILIC, showed a slightly shorter arrival time. Only a single peak of very low intensity was observed to tetragallic acid (**IV**) in both HILIC and RP-LC.

4.4 Conclusions

Comprehensive analysis of a commercial tara extract by RP-LC and HILIC coupled to diode array detection and IM-HR-MS allowed for the tentative identification of 43 isomeric galloyl-quinic acid species in addition to six gallic acid derivatives and ellagic acid. RP-LC provided superior separation for tara gallotannins compared to HILIC. The complimentary information provided by the combination of UV spectral data and IM separation provided a facile methodology to identify the number of depsidic bonds, and thereby, to differentiate between positional isomers of the same degree of galloylation. Characteristic arrival time differences allowed in most cases for the assignment of individual isomeric species between HILIC and RP-LC. The proposed presence of prototropic isomers explains the loss in mass spectrometric specificity for the mono-galloylquinic acid species as well as multiple arrival times for the di-galloyl-quinic acid isomers. Evidence of on-column conversion between species was observed in RP-LC, where the combination of IM and chromatographic data suggest the migration of galloyl groups to different positions on the quinic acid core as well as to galloyl phenolic groups.

The results of the present study for tara gallotannins, combined with the findings reported in the first part of this work on ellagitannins in chestnut (chapter 3), confirm the utility of an analytical methodology based on the combination of RP-LC and HILIC separation with IM and HR-MS for the comprehensive analysis of hydrolysable tannins.

4.5 References

1. Haslam E, Haworth RD, Keen PC (1962) Gallotannins. Part VII. Tara Gallotannin. J Chem Soc 3814–3818

2. Armitage R, Bayliss GS, Gramshaw JW, Haslam E, Haworth RD, Jones K, Rogers HJ, Searle T (1961) Gallotannins. Part III.* The Constitution of Chinese, Turkish, Sumach, and Tara Tannins. J Chem Soc, Perkin Trans 2 1842–1853

3. Arapitsas P, Menichetti S, Vincieri FF, Romani A (2007) Hydrolyzable tannins with the hexahydroxydiphenoyl unit and the m-depsidic link: HPLC-DAD-MS identification and model synthesis. J Agric Food Chem 55:48–55

4. Martins D, Duarte L, Silva VFM, Crispim A, Beghini E, Crispim F (2018) Study of vegetable extracts effect on wet-white leather. Leather Footwear J 18:213–218

5. Lagel MC, Pizzi A, Giovando S (2014) Matrix-Assisted Laser Desorption-Ionization Time of Flight (MALDI-TOF) Mass Spectrometry of Phenol-Formaldehyde-Chestnut Tannin Resins. J Renew Mater 2:207–219

6. Aouf C, Benyahya S, Esnouf A, Caillol S, Boutevin B, Fulcrand H (2014) Tara tannins as phenolic precursors of thermosetting epoxy resins. Eur Polym J 55:186–198

7. Galvez JMG, Riedl B, Conner AH (1997) Analytical Studies on Tara Tannins. Holzforschung 51:235–243

8. Zhao B, Han W, Zhang W, Shi B (2018) Corrosion inhibition performance of tannins for mild steel in hydrochloric acid solution. Res Chem Intermed 44:407–423

9. Merino SF, Caprari JJ, Torres LV, Ramos LF, Girola AH (2017) Inhibitive action of tara tannin in rust converter formultaion. Anti-Corrosion Methods Mater 64:136–147

10. Aguilar-Galvez A, Noratto G, Chambi F, Debaste F, Campos D (2014) Potential of tara (Caesalpinia spinosa) gallotannins and hydrolysates as natural

antibacterial compounds. Food Chem 156:301–304

11. Tian F, Li B, Ji B, Zhang G, Luo Y (2009) Identification and structure-activity relationship of gallotannins separated from Galla chinensis. LWT - Food Sci Technol 42:1289–1295. doi: 10.1016/j.lwt.2009.03.004

12. Arapitsas P (2012) Hydrolyzable tannin analysis in food. Food Chem 135:1708–1717. doi: 10.1016/j.foodchem.2012.05.096

13. Mueller-Harvey A (2001) Analysis of hydrolysable tannins. Anim Feed Sci Technol 91:3–20

14. Engels C, Knödler M, Zhao YY, Carle R, Gänzle MG, Schieber A (2009) Antimicrobial activity of gallotannins isolated from mango (Mangifera indica L.) kernels. J Agric Food Chem 57:7712–7718. doi: 10.1021/jf901621m

15. Salminen JP, Ossipov V, Loponen J, Haukioja E, Pihlaja K (1999) Characterisation of hydrolysable tannins from leaves of Betula pubescens by high-performance liquid chromatography-mass spectrometry. J Chromatogr A 864:283–291. doi: 10.1016/S0021-9673(99)01036-5

16. Owen RW, Haubner R, Hull WE, Erben G, Spiegelhalder B, Bartsch H, Haber B (2003) Isolation and structure elucidation of the major individual polyphenols in carob fibre. Food Chem Toxicol 41:1727–1738. doi: 10.1016/S0278-6915(03)00200-X

17. Delahaye P, Verzele M (1983) Analysis of gallic, digallic and trigallic acids in tannic acids by high-performance liquid chromatography. J Chromatogr 265:363–367

18. Beasley TH, Ziegler HW, Bell AD (1991) Determination and Characterization of Gallotannin by High Performance Liquid Chromatography. Anal Chem 49:238–243. doi: 10.1002/bmc.1130050503

19. Clifford MN, Stoupi S, Kuhnert N (2007) Profiling and Characterization by LC-MS of the Galloylquinic Acids of Green Tea, Tara Tannin, and Tannic Acid. J Agric Food Chem 55:2797–2807

20. Clifford MN, Johnston KL, Knight S, Kuhnert N (2003) Hierarchical Scheme for LC-MSn Identification of Chlorogenic Acids. J Agric Food Chem 51:2900–2911

21. Clifford MN, Knight S, Kuhnert N (2005) Discriminating between the six isomers of dicaffeoylquinic acid by LC-MSn. J Agric Food Chem 53:3821–3832

22. Clifford MN (1986) Coffee bean dicaffeoylquinic acids. Phytochemistry 25:1767–1769

23. Clifford MN, Knight S, Birgul S, Kuhnert N (2006) Characterization by LC-MSn of Four New Classes of Chlorogenic Acids in Green Coffee Beans: Dimethoxycinnamoylquinic Acids, Diferuloylquinic Acids, Caffeoyl-dimethoxycinnamoylquinic Acids, and Feruloyl-dimethoxycinnamoylquinic Acids. J Agric Food Chem 54:1957–1969

24. Mane C, Sommerer N, Yalcin T, Cheynier V, Cole RB, Fulcrand H (2007) Assessment of the Molecular Weight Distribution of Tannin Fractions through MALDI-TOF MS Analysis of Protein−Tannin Complexes Assessment of the Molecular Weight Distribution of Tannin Fractions through MALDI-TOF MS Analysis of Protein-Tannin Complexes. Anal Chem 79:2239–2248

25. Pizzi, A. Pasch, H. Rode, K. Giovando S (2009) Polymer Structure of Commercial Hydrolyzable Tanins by Matrix-Assisted Laser Desorption/Ionization-Time-of-Flight Mass Spectriometry. J Appl Polym Sci 113:3847–3859

26. Franceschi P, Vrhovsek U, Guella G (2011) Ion mobility mass spectrometric investigation of ellagitannins and their non-covalent aggregates. Rapid Commun Mass Spectrom 25:827–833

27. Xie C, Yu K, Zhong D, Yuan T, Ye F, Jarrell JA, Millar A, Chen X (2011) Investigation of Isomeric Transformations of Chlorogenic Acid in Buffers and Biological Matrixes by Ultraperformance Liquid Chromatography Coupled with Hybrid Quadrupole/Ion Mobility/Orthogonal Acceleration Time-of-Flight Mass Spectrometry. J Agric Food Chem 59:11078–11087

28. Kuhnert N, Yassin GH, Jaiswal R, Matei MF, Grün CH (2015) Differentiation of

prototropic ions in regioisomeric caffeoyl quinic acids by electrospray ion mobility mass spectrometry. Rapid Commun Mass Spectrom 29:675–680

29. Zheng X, Renslow RS, Makola MM, Webb IK, Deng L, Thomas DG, Govind N, Ibrahim YM, Kabanda MM, Dubery IA, Heyman HM, Smith RD, Madala NE, Baker ES (2017) Structural Elucidation of cis/trans Dicaffeoylquinic Acid Photoisomerization Using Ion Mobility Spectrometry-Mass Spectrometry. J Phys Chemistry Lett 8:1381–1388

30. Ruotolo BT, Benesch JL, Sandercock AM, Hyung S, Robinson C V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes. Nat Protoc 3:1139–1152

31. Forsythe JG, Petrov AS, Walker CA, Allen SJ, Pellissier JS, Bush MF, Hud N V., Fernández FM (2015) Collision cross section calibrants for negative ion mode traveling wave ion mobility-mass spectrometry. Analyst 140:6853–6861

32. Donovan JL, Meyer AS, Waterhouse A (1998) Phenolic Composition and Antioxidant Activity of Prunes and Prune Juice (Prunus domestica) Prune Juice (Prunus domestica). J Agric Food Chem 46:1247–1252

33. Nakatani N, Kayano S, Kikuzaki H, Sumino K, Katagiri K (2000) Identification, Quantitative Determination, and Antioxidative Activities of Chlorogenic Acid Isomers in Prune (Prunus domestica L). J Agric Food Chem 48:5512–5516

34. Fang N, Yu S, Prior RL (2002) LC/MS/MS characterization of phenolic constituents in dried plums. J Agric Food Chem 50:3579–3585

35. Warnke S, Seo J, Boschmans J, Sobott F, Scrivens JH, Bleiholder C, Bowers MT, Gewinner S, Schöllkopf W, Pagel K, Von Helden G (2015) Protomers of benzocaine: Solvent and permittivity dependence. J Am Chem Soc 137:4236–4242. doi: 10.1021/jacs.5b01338

36. Gagliardi LG, Castells CB, Ràfols C, Rosés M, Bosch E (2007) δ conversion parameter between pH scales (WSpH and SSPH) in acetonitrile/water mixtures at various compositions and temperatures. Anal Chem 79:3180–3187. doi: 10.1021/ac062372h

37. Willemse CM, Stander MA, de Villiers A (2013) Hydrophilic interaction chromatographic analysis of anthocyanins. J Chromatogr A 1319:127–140

38. Bu X, Skrdla PJ, Dormer PG, Bereznitski Y (2010) Separation of triphenyl atropisomers of a pharmaceutical compound on a novel mixed mode stationary phase: A case study involving dynamic chromatography, dynamic NMR and molecular modeling. J Chromatogr A 1217:7255–7264

39. Fedurcová A, Vančová M, Mydlová J, Lehotay J, Krupčík J, Armstrong DW (2006) Interconversion of Oxazepam Enantiomers During HPLC Separation. Determination of Thermodynamic Parameters. J Liq Chromatogr Relat Technol 29:2889–2900

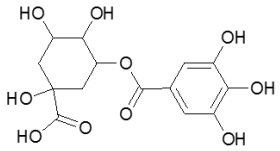
40. Verzele M, Delahaye P, van Dijck J (1983) Digallic Acid. Bull des Sociétés Chim Belges 92:181–186. doi: 10.1002/bscb.19830920212

41. Nierenstein M, Spiers CW, Hatcher PR (1925) Gallotannin. XIII. The identity of digallic acid from gallotannins with synthetic meta-digallic acid. J Am Chem Soc 47:846–850

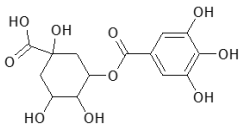
Chapter 4

Supporting information for: Comprehensive analysis of hydrolysable tannins by reversed phase and hydrophilic interaction chromatography coupled to ion mobility and high-resolution mass spectrometry, Part 2: Tara tannins

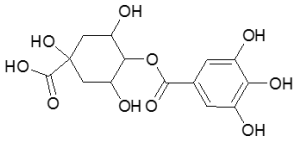
3-O-galloylquinic acid (**2**[a])



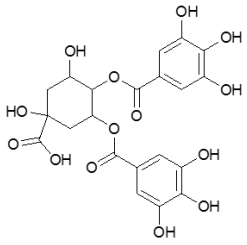
5-O-galloylquinic acid (**2**[b])



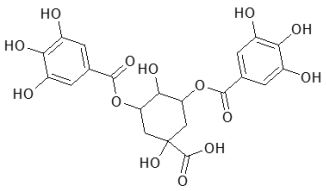
4-O-galloylquinic acid (**2**[c])



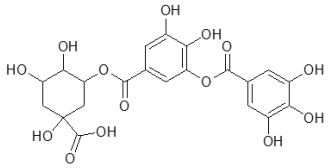
3,4-digalloylquinic acid (**2**[a])



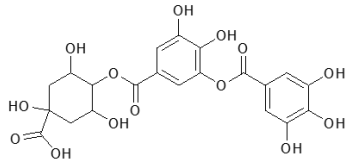
3,5-digalloylquinic acid (**2**[b])†



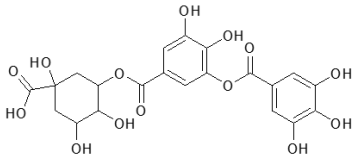
3-O-digalloylquinic acid (**2**[c])†



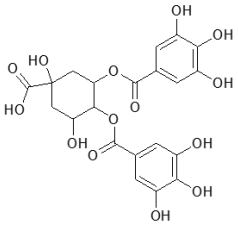
5-O-digalloylquinic acid (**2**[f])†



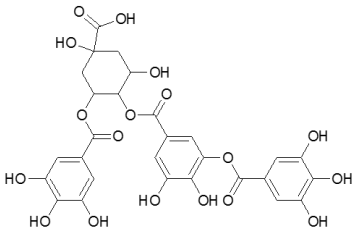
5-O-digalloylquinic acid (**2**[d])†



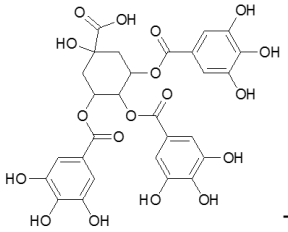
4,5-digalloylquinic acid (**2**[e])



Tri-O-galloylquinic acid (2 depsidic bond) (**3**[b-h])*†

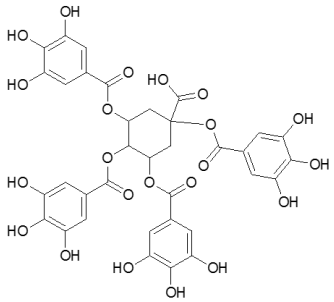


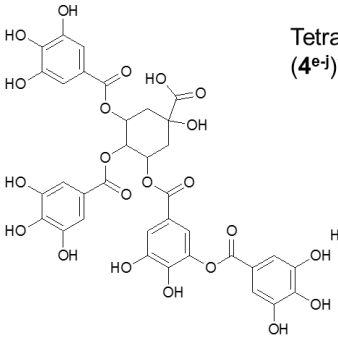
3,4,5-tri-O-galloylquinic acid (**3**[a])



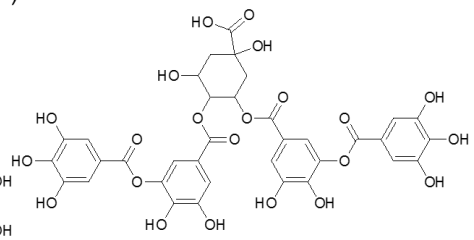
Tetra-O-galloylquinic acid (2 depsidic bond)

2,3,4,5-tetra-O-galloylquinic acid (**4**[a]) (**4**[b-d])*†

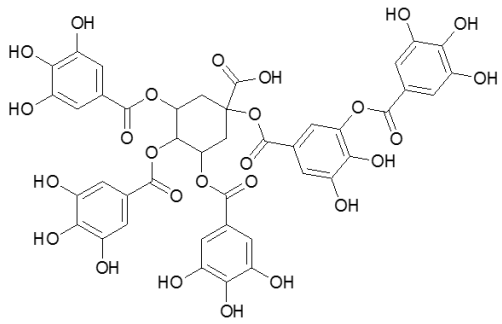




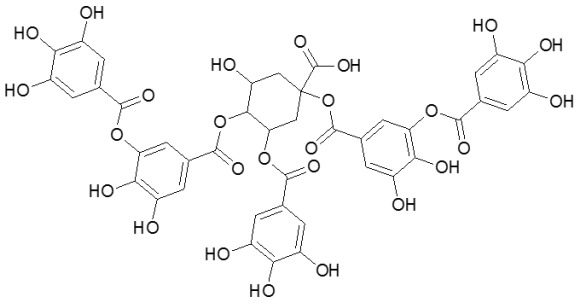
Tetra-O-galloylquinic acid (2 depsidic bonds) (**4**[e-j])*†



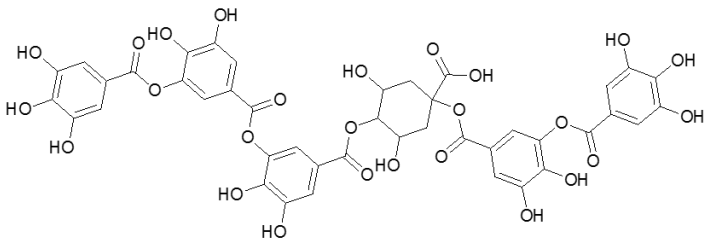
2-O-di-3,4,5-tri-O-galloylquinic acid (**5**[b])[#†]



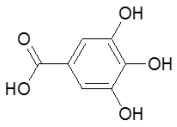
Penta-O-galloylquinic acid (2 depsidic bonds) (**5**[a,c-f,m])[*†]



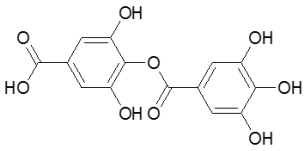
Penta-O-galloylquinic acid (3 depsidic bonds) (**5**[g-o])[*†]



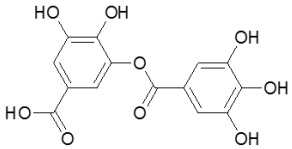
Gallic acid (**I**)



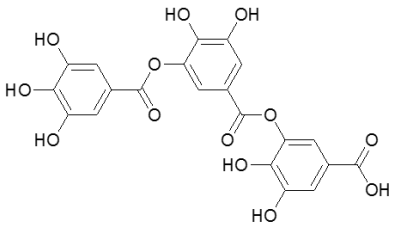
p-Digallic acid (**II**[a])



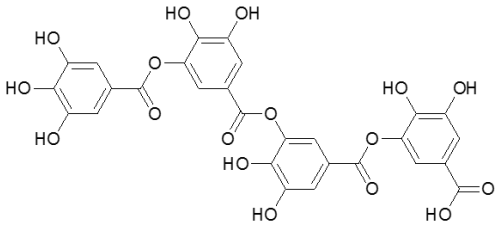
m-Digallic acid (**II**[b])



Trigallic acid (**III**)[†]



Tetragallic acid (**IV**)[†]



Ellagic acid (**V**)

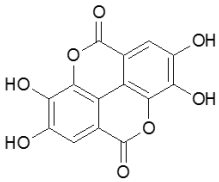


Figure S1: Structures of the galloyl-quinic acid derivatives tentatively identified in the tara extract. Compound marked with [#] are tentatively identified for the first time in tara. For compounds labelled *, representative structures are shown, since regioisomers could not be identified based on the available data. Furthermore, for compounds labelled †, only representative structures are presented, as *m*- and *p*- isomers could not be distinguished.

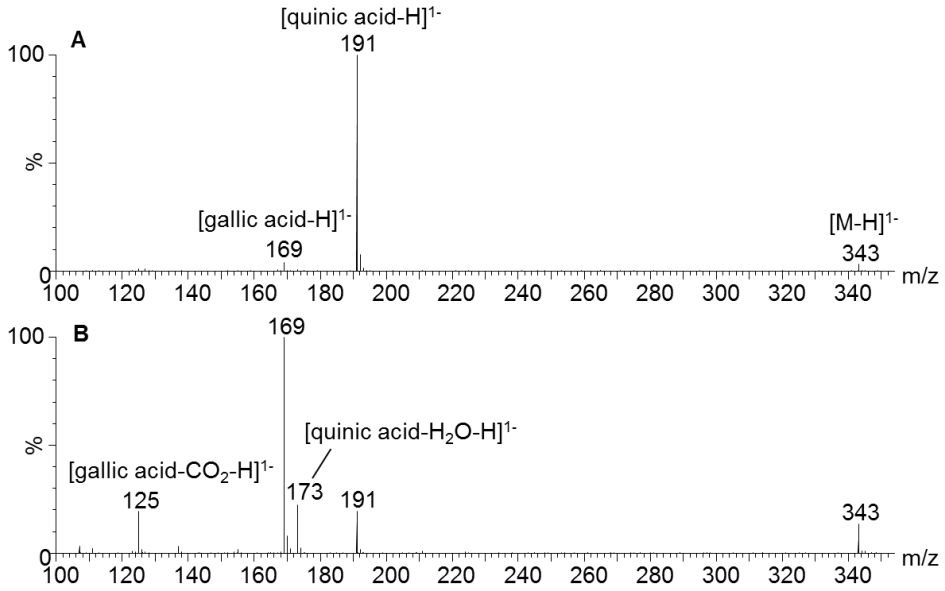


Figure S2: MS^E spectra of (**A**) 5-O-galloylquinic acid (**1[b]**) and (**B**) 4-O-galloylquinic acid (**1[c]**) eluting at 2.28 and 2.80 min, respectively, under RP-LC conditions.

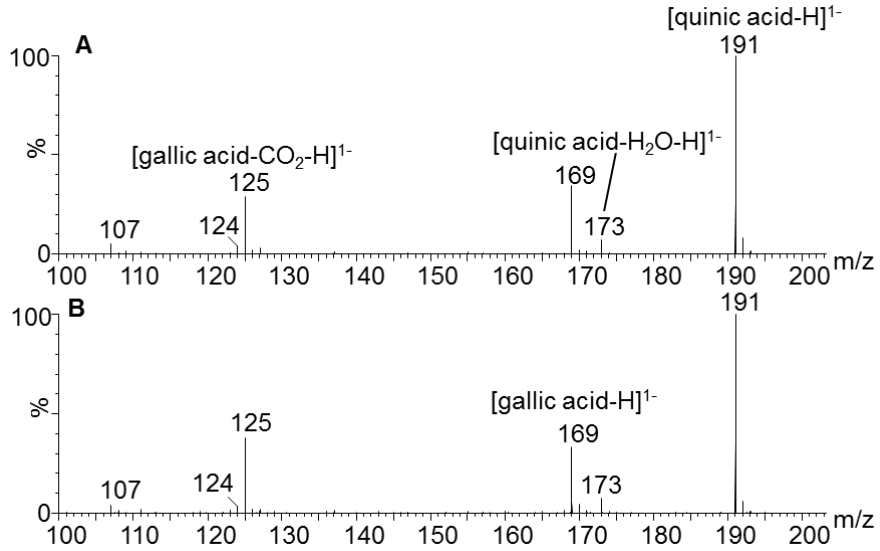


Figure S3: MS^E spectra obtained for the two major monogalloylquinic acid (**1**) isomers eluting at 6.59 and 7.62 minutes, respectively, in HILIC.

.

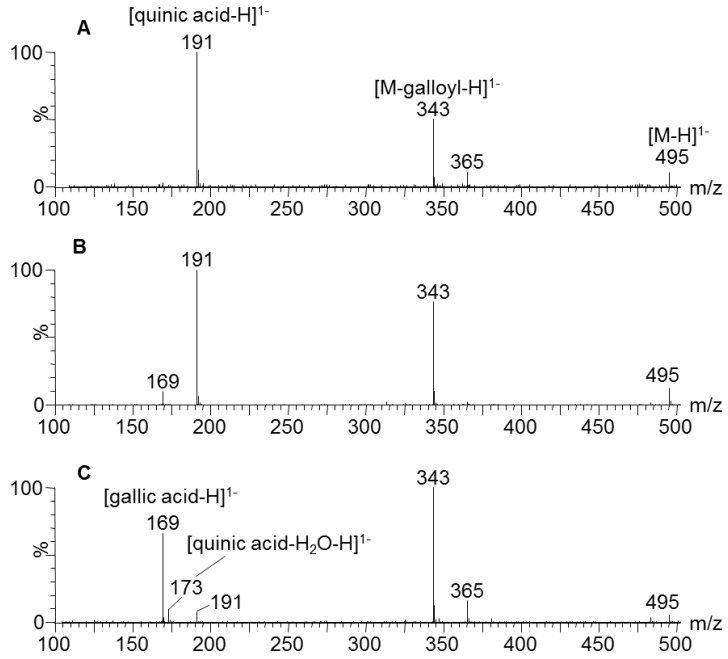


Figure S4: MS^E spectra obtained for (**A**) 3-O-digalloyl quinic acid (**2^c**), (**B**) 5-O-digalloyl quinic acid (**2^d**), and (**C**) 4-O-digalloyl quinic acid (**2^f**) under RP-LC conditions.

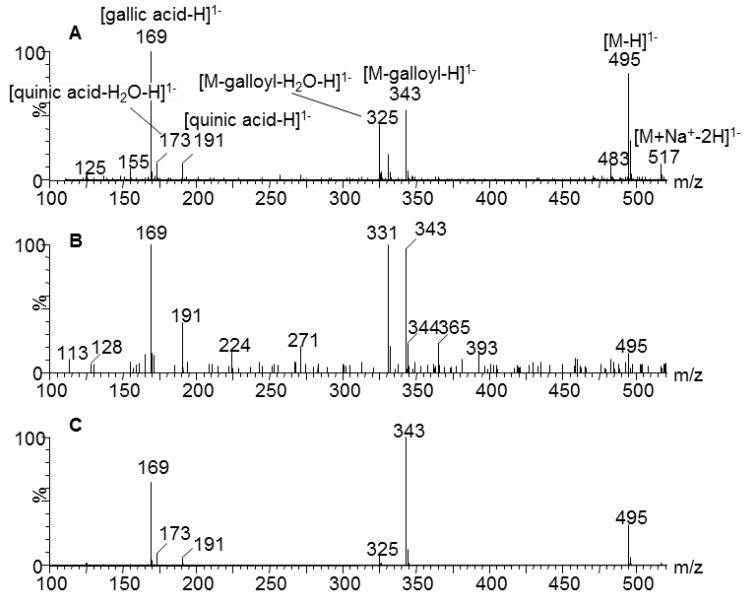


Figure S5: MS^E spectra obtained for (**A**) 3,4- di-O-galloyl quinic acid (**2^a**), (**B**) 3,5- di-O-galloyl quinic acid (**2^b**), and (**C**) 4,5- di-O-galloyl quinic acid (**2^e**) under RP-LC conditions.

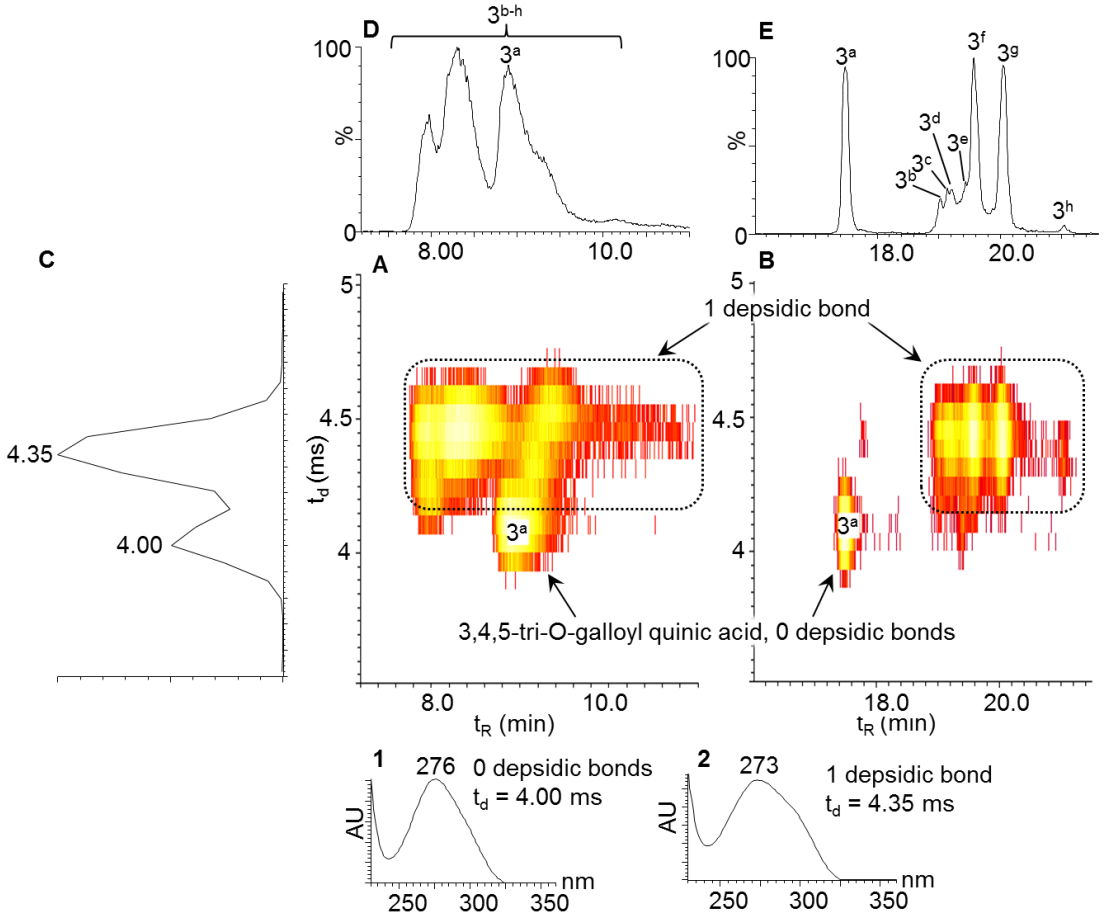


Figure S6: HILIC×ion mobility (**A**) and RP-LC×ion mobility (**B**) contour plots showing the correlation between arrival times and UV spectra for the trigalloyl-quinic acid species. Peak labels correspond to **Table 4.1**. (**C**) shows the extracted ion arrival plot for m/z 647 (identical for RP-LC and HILIC), while (**D**) and (**E**) show the corresponding extracted ion chromatograms for HILIC and RP-LC, respectively.

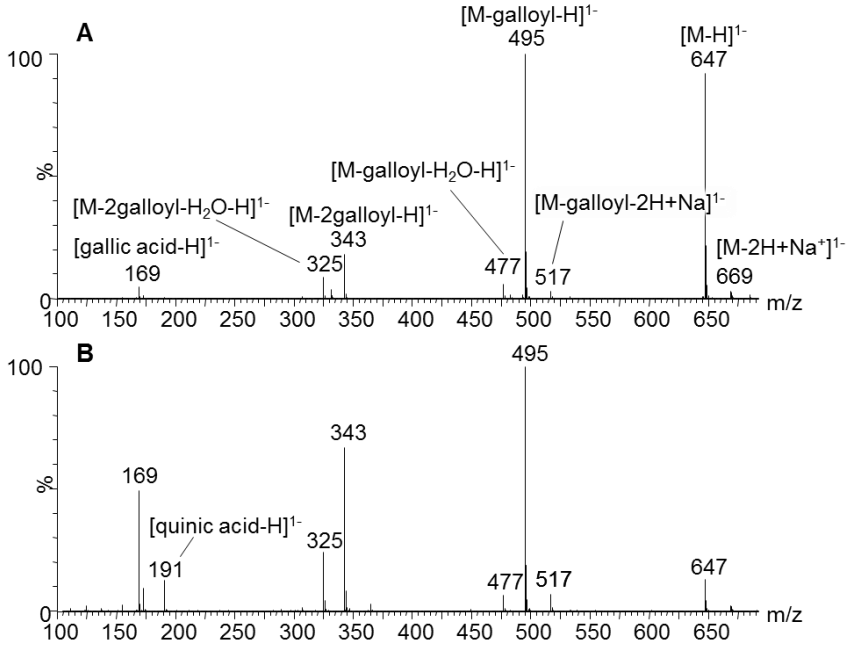
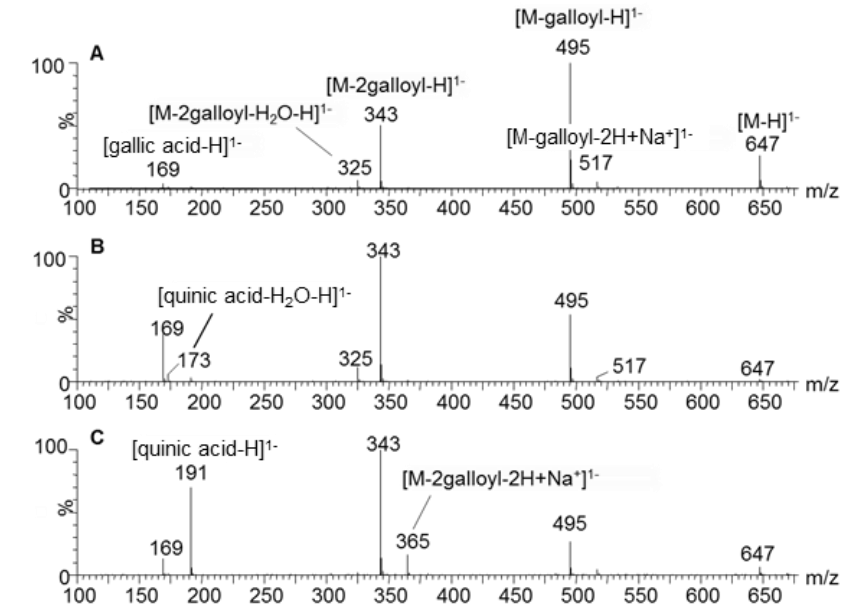


Figure S7: MS^E spectra of 3,4,5-tri-O-galloyl-quinic acid (**3ª**) obtained under RP-LC (**A**) and HILIC (**B**) conditions.



.

Figure S8: (**A**) MS^E spectra of trigalloyl-quinic acid isomers **3b-h** obtained by RP-LC (identical spectra were observed for all isomers). (**B**) and (**C**) show MS^E spectra obtained in HILIC for the trigalloyl-quinic acid isomers eluting at 8.30 (**B**) and 9.30 (**C**) min, respectively.

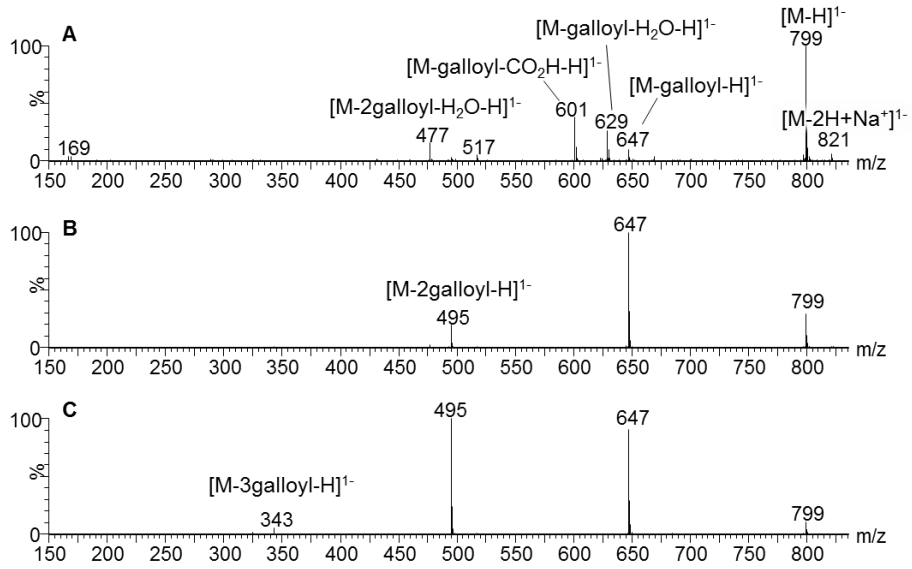


Figure S9: MS^{E} spectra obtained for tetragalloyl-quinic acid isomers with arrival times of (**A**) 4.76 ms (**4^{a}**, no depsidic bonds), (**B**) 5.04, 5.11 and 5.24 ms (**4^{b-d}**, one depsidic bond, all identical), and (**C**) 5.66 ms (**4^{e-j}**, two depsidic bonds, all identical) under RP-LC conditions.

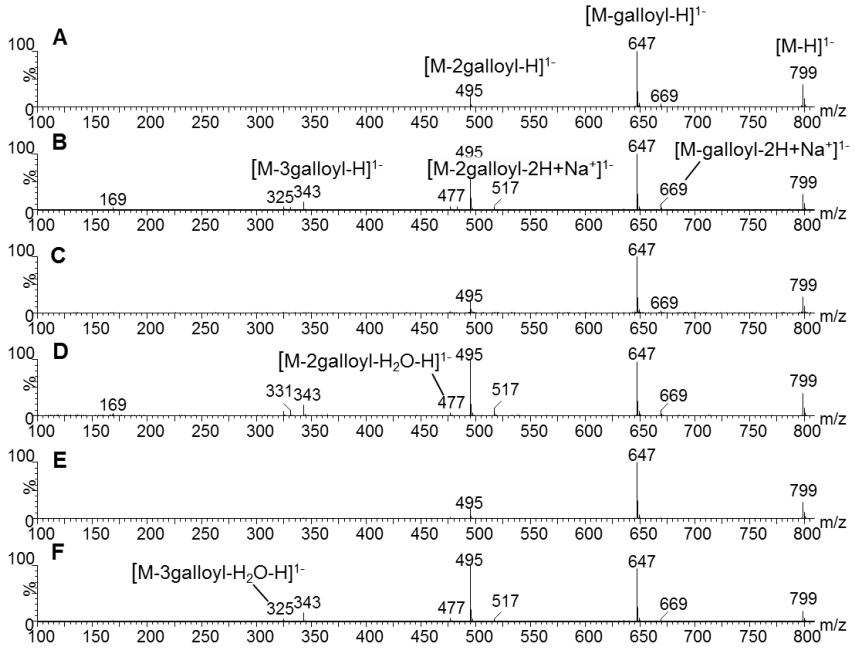


Figure S10: MS^{E} spectra of tetragalloyl-quinic acid species containing one depsidic bond (**4^{b}**, **4^{c}** and **4^{d}**) at arrival times of 5.04 (**A, B**), 5.11 (**C, D**) and 5.24 (**E, F**) ms, respectively, under RP-LC (**A, C** and **E**) and HILIC (**B, D** and **F**) conditions.

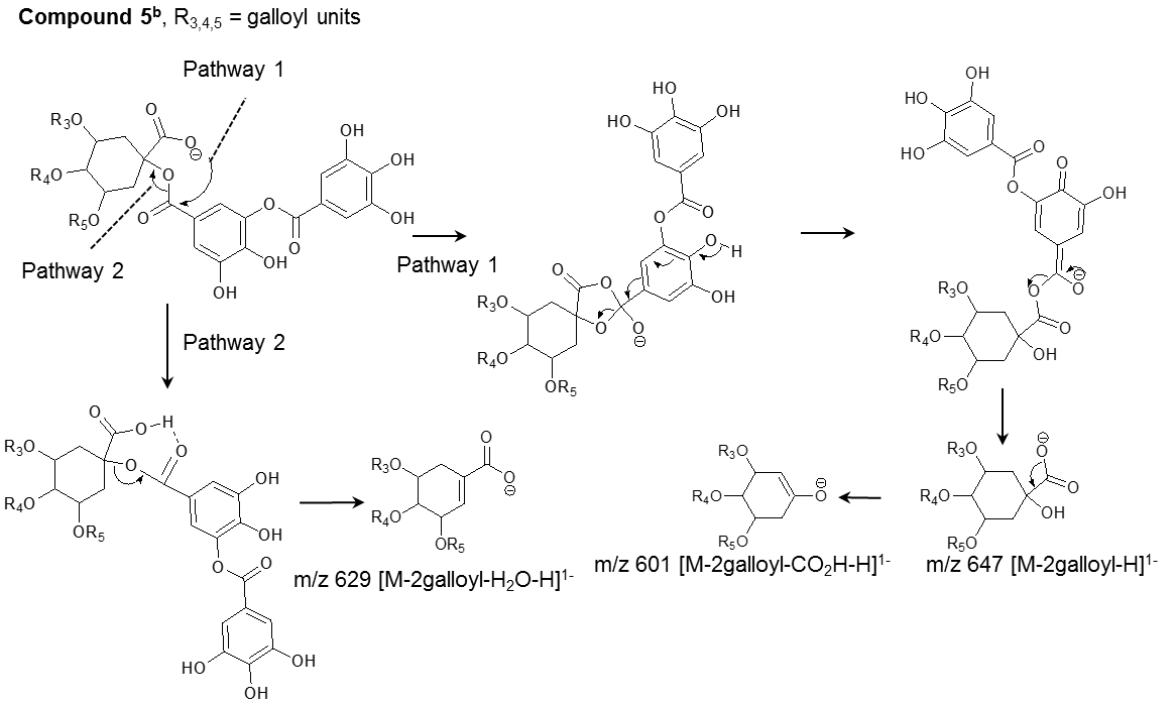


Figure S11: Proposed fragmentation of 1-O-digalloyl-3,4,5-tri-O-galloylquinic acid (**5^{b}**) indicating the formation of the m/z 629, 647 and 601 fragment ions. Adapted from scheme B in [21].

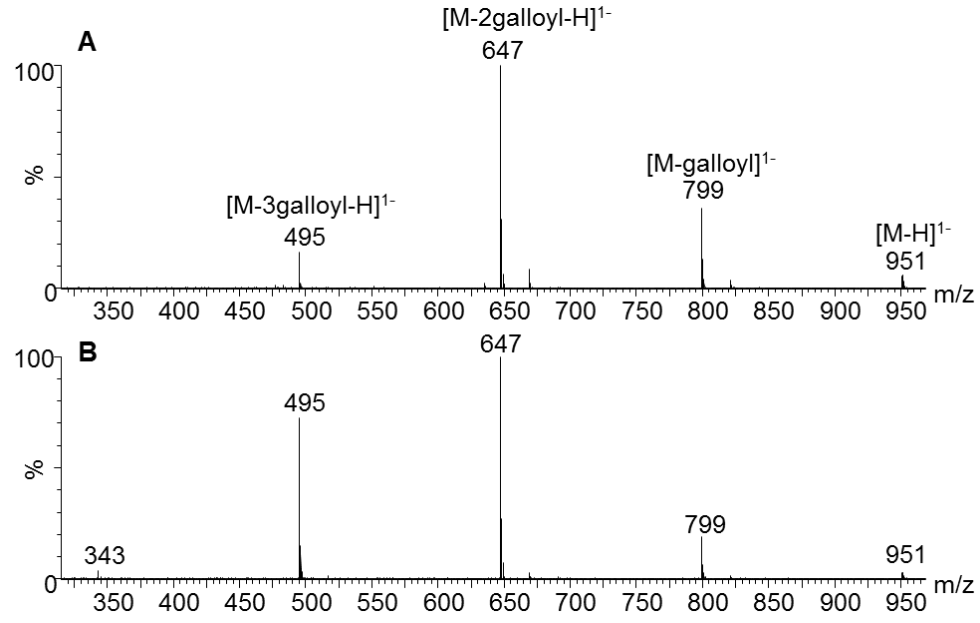


Figure S12: Representative MS^{E} fragmentation spectra of pentagalloyl-quinic acid isomers with arrival times of 6.28 (**A**, **5$^{a,c-f,m}$**) and 7.04 (**B, 5^{g-o}**) ms.

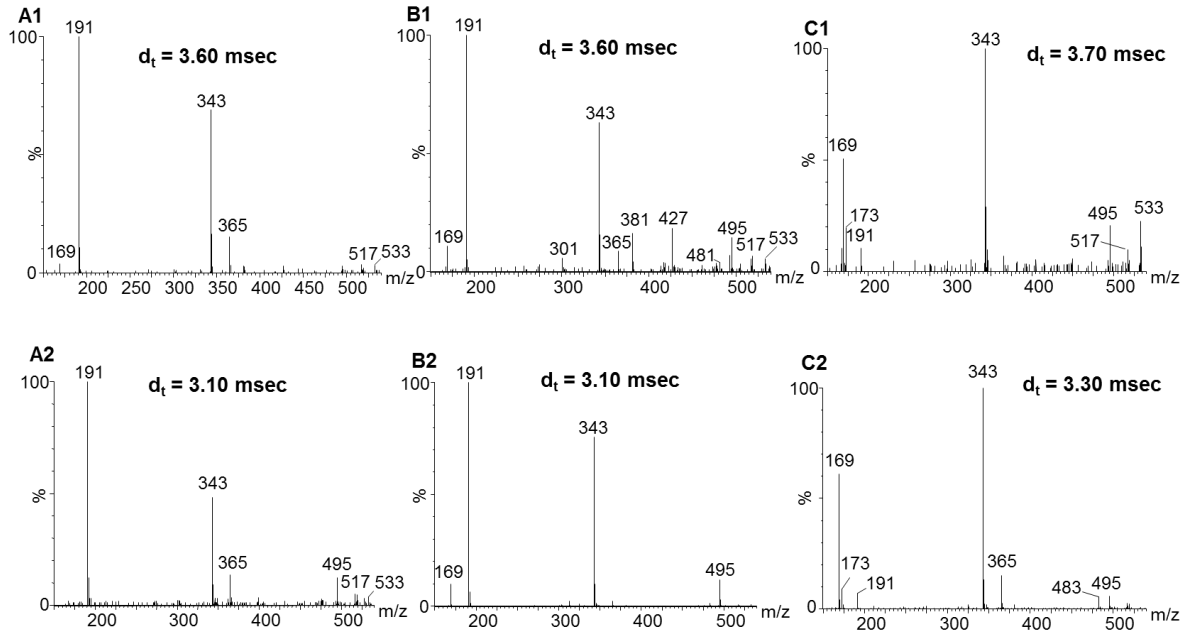


Figure S13: Arrival time filtered MS^E spectra obtained for the two IM resolved species of 3-O-digalloyl quinic acid (**2^c**, **A1,** 3.60 and **A2,** 3.10 ms), 5-O-digalloyl quinic (**2^d**, **B1,** 3.60 and **B2,** 3.10 ms), and 4-O-digalloyl quinic (**2^f**, **C1,** 3.70 and **C2,** 3.30 ms).

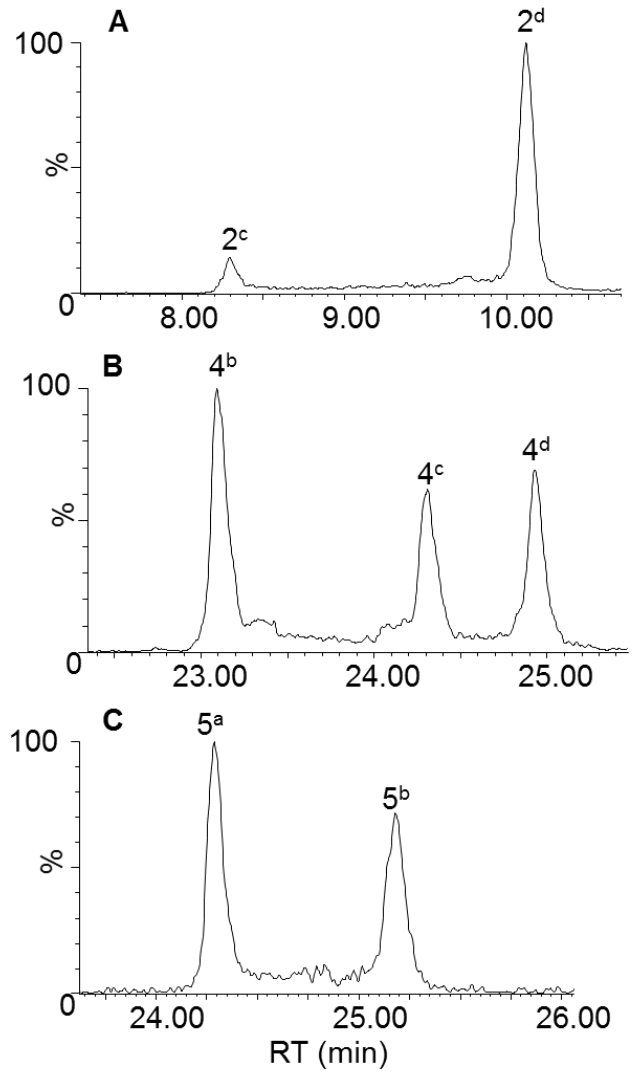


Figure S14: Examples of on-column inter-conversion between di- (**2**^{c,d}**, A**), tetra- (**4**^{b,c,d}**, B**) and penta- (**5**^{a,b}**, C**) galloyl-quinic acid derivatives in tara under RP-LC conditions. Extracted ion chromatograms are presented for m/z 495 (**2**^{c,d}, **A**), 799 m/z (**4**^{b,c,d}, **B**) and 951 m/z (**5**^{a,b}, **C**). Note the baseline offset between individual isomers, which indicates on-column conversion between these species.

Chapter 5

On-line comprehensive two-dimensional hydrophilic interaction chromatography × reversed phase liquid chromatography - mass spectrometric analysis of hydrolysable tannins

Abstract

Extending on the work reported in Chapters 3 and 4, this chapter reports the development and implementation of on-line comprehensive two-dimensional hydrophilic interaction chromatography × reversed-phase liquid chromatography (HILIC×RP-LC) methods hyphenated to high resolution mass spectrometric (HR-MS) detection for the analysis of hydrolysable tannins. Industrial chestnut and tara tannin samples, comprising respectively mainly ellagi- and gallo-tannins, respectively, were analysed. HILIC×RP-LC methods were optimised using a predictive kinetic optimisation protocol, where the first dimension (^{1}D) effluent was diluted with an aqueous make-up flow to minimise second dimension (^{2}D) band broadening effects. The optimised methods provided practical peak capacities in excess of a thousand, which confirms their utility for the high-resolution separation of hydrolysable tannins. The combination of a narrow-bore ^{1}D column and a high ^{2}D flow rate dictated by sampling time constraints resulted in severe dilution, and negatively impacted on the overall sensitivity and resolution of the HILIC×RP-LC methods. Nevertheless, group type separation was observed for both chestnut and tara hydrolysable tannins, which facilitated tentative assignment of compound classes and may prove beneficial for screening analysis of hydrolysable tannins. To our knowledge this is the first report of the comprehensive two-dimensional LC analysis of hydrolysable tannins.

5.1 Introduction

Hydrolysable tannins include both gallotannin and ellagitannin compound classes, which are distinguished by the formation of gallic and ellagic acid, respectively, following hydrolysis [1, 2]. Tara (*Caesalpinia spinosa*) contains gallotannins comprised of a quinic acid polyol core connected to gallic acid moieties through ester linkages [2]. In contrast, chestnut (*Castanea sativa*) tannins include both gallotannins and ellagitannins, the latter containing an open or closed glucose core esterified to hexahydroxydiphenic acid (HHDP) or nonahydroxyterphenoyl (NHTP) groups and/or gallic acid moieties [3].

The analysis of hydrolysable tannins is challenging due to their complexity: both ellagitannins and gallotannins typically comprise large numbers of isomeric compounds. In Chapter 3 it was demonstrated that one dimensional (1D) hydrophilic interaction (HILIC) and reversed phase (RP-LC) analyses provide complimentary information for chestnut tannins, with superior separation for particular compound classes obtained by either separation mode. For example, sixteen isomeric species of vescalagin/castalagin digalloyl glucose were resolved by RP-LC compared to only one peak detected under HILIC conditions. RP-LC also showed superior selectivity for the galloyl glucose species, whereas no co-elution of vescalagin, castalagin and their dimeric counterparts, roburins A-D were observed under HILIC conditions. On the other hand, RP-LC showed much better separation of tara gallotannins, with relatively poor chromatographic performance for these compounds obtained on an Amide HILIC column (chapter 4). Furthermore, a loss of mass spectrometric specificity was noted for some tara gallotannins (chapter 4) under HILIC conditions (i.e. high acetonitrile concentrations).

Considering the above, the comprehensive combination of HILIC and RP-LC may therefore be a promising approach for the detailed analysis of hydrolysable tannins. Indeed, comprehensive two-dimensional HILIC×RP-LC has found extensive application in the analysis of condensed tannins in food and natural products [4–8]. To the best of our knowledge, the use of LC×LC has not been reported for hydrolysable tannins to date. The goal of the work reported in the present chapter was therefore to develop and implement on-line HILIC×RP-LC methods for the comprehensive analysis of chestnut and tara hydrolysable tannins. The developed LC×LC methods will be

hyphenated to HR-MS to allow compound identification and to demonstrate the general applicability of this approach for hydrolysable tannin analysis.

5.2 Experimental

5.2.1 Materials and reagents

The tara sample were donated by Silvateam (San Michele Mondovì (Cuneo), whereas the chestnut sample is produced by extracting wood chips with 2% (m/v) $NaHSO_3$ and 0.5% $NaHCO_3$ in hot water (70-80°C). HPLC grade methanol (MeOH) and formic acid (FA) were purchased from Sigma-Aldrich (Johannesburg, South Africa) and HPLC grade acetonitrile (ACN) from ROMIL (Waterbeach Cambridge, England). Deionised water was obtained using a Milli-Q water purification system (Millipore, Milford, MA).

5.2.2 Sample preparation

Chestnut (49 mg) and tara (12 mg) samples were dissolved in 1 mL MeOH/H_2O (50:50, (v/v)) and MeOH/ACN (10:90, (v/v)), respectively. Samples were filtered through 0.45 µm hydrophilic PVDF membrane filters (Millipore) before analyses.

5.2.3 Instrumentation

For the analysis of both samples, a Waters Capillary LC 920 pump and autosampler equipped with a 1 µL loop and controlled by MassLynx v4.1 software (Waters, Milford, MA) was used in the first dimension (^{1}D). In the second dimension (^{2}D) an Agilent 1290 Infinity binary pump connected via a 2-position/8-port valve equipped with two 80 µL loops controlled by OpenLab CDS software (Agilent Technologies, Waldbronn, Germany) was used. To provide a make-up flow of pure water, an Agilent 1100 Isocratic pump was connected via a T-piece between the ^{1}D column and the 2-position/8-port valve. An Agilent 1200 column oven was used to thermostat the ^{2}D column at 50°C. A 1:4 flow split was installed between the Agilent 1290 Infinity II DAD detector (1 µL flow cell, 190-640 nm, 80 Hz) and a Waters Synapt G2 quadrupole-time-of-flight (Q-TOF) mass spectrometer (100-2000 amu, 5 Hz). The mass spectrometer was equipped with an electrospray ionisation source operating in the negative ion mode with capillary and cone voltages of -2.5 kV and 15 V, respectively. A source temperature of 120°C was used, combined with a desolvating temperature

and gas flow of 275°Cand 650 L/h, respectively. Low (4eV) and high (ramped from 10 to 30 eV) collisional spectra were acquired concurrently by operating the instrument in MS^E mode. Mass spectral data were acquired using a mass range of 100-2000 amu with a scan time of 0.2 seconds. Nitrogen was used as the ion mobility drift gas at a flow of 90 mL/min and helium with a flow of 180 mL/min in the helium cell. Ion mobility separations were accomplished using a wave velocity of 448 m/s with a wave height of 37.1 V and a transfer velocity of 380 m/s.

5.2.4 Chromatographic conditions

For ^{1}D HILIC separations, an Acquity UPLC BEH Amide column (150 × 1 mm i.d., 1.7 µm, Waters) was used at room temperature (25°C). ^{2}D RP-LC separations were performed on a Kinetex C_{18} core-shell column (50 × 3.0 mm i.d., 1.7 µm, Phenomenex, Torrance, USA) at 50°C. An injection volume of 1 µL was used for both samples. Flow rates of 9 µL/min and 3 mL/min were used in the first and second dimensions, respectively. For both samples and dimensions mobile phase A consisted of 0.1% FA (v/v) and mobile phase B of ACN. Further chromatographic conditions, including make-up flows, modulation periods and gradient times are summarised in **Table 5.1**.

Table 5.1: Chromatographic conditions used for the on-line comprehensive HILIC×RP-LC separation of tara and chestnut tannins.

Sample	**Tara**		**Chestnut**	
Dimension	**^{1}D**	**^{2}D**	**^{1}D**	**^{2}D**
Gradient	10-20% A in 0-50 min	1-35%B in 0-0.27 min; 35-100% B in 0.27-0.28 min	10-20% A in 0-20 min;20-50% A in 20-30 min; 50-98% A in 35-50 min; 98% A 50-60 min	1-15%B in 0-0.15 min; 15-45% B in 0.15-0.27 min; 45-100% in 0.27-0.28 min
Dilution flow (µL/min)		126		126
Modulation time (min)		0.47		0.47
Gradient time (min)		0.27		0.27

5.3 Results and discussion

5.3.1 On-line HILIC×RP-LC separation conditions

The current study follows on previous work done on the comprehensive analysis of chestnut and tara hydrolysable tannins using both 1D RP-LC-IM-MS and HILIC-IM-MS (Chapters 3 and 4). In the present chapter, a comprehensive on-line HILIC×RP-LC method for the analysis of these samples was developed.

During an on-line comprehensive LC×LC separation, the ^{1}D effluent is collected (sampled) while simultaneously analysing the previously collected fraction on the ^{2}D column. This is typically done using a switching valve equipped with two sampling loops (**Figure 5.1**). Since both separations are carried out concurrently, the ^{1}D sampling time is equal to the ^{2}D cycle time. Furthermore, to prevent ^{1}D under-sampling, each ^{1}D peak should be sampled 3 times [9], which implies very short modulation periods and ^{2}D analysis times. Since HILIC eluents are strong solvents in RP-LC, and fractions are directly transferred to the ^{2}D column in on-line LC×LC, it is vital to minimise fraction volumes as far as possible to avoid excessive band broadening in the ^{2}D. The HILIC×RP-LC conditions used in this work were derived using a predictive kinetic optimisation approach previously reported [10]. This procedure included the use of fixed active modulation, where injection band broadening is avoided by limiting the fraction volume and diluting the first dimension eluent with a weak ^{2}D solvent before the modulation valve (**Figure 5.1,** orange arrow). The columns used in the optimisation consisted of a 1 mm i.d. amide column in the first dimension, combined with a C_{18} 1.7 µm core-shell column in the second dimension. Experimental conditions optimised using the multi-parameter optimisation protocol include the flow rates used in both dimensions, the sampling time, dilution flow and ^{2}D gradient time. In order to achieve optimal ionisation, a 1:4 split was installed before the mass spectrometer. The experimental conditions used are summarised in **Table 5.1**. 2D UV contour plots obtained for the optimised conditions are presented in **Figures 5.2** and **5.5** for chestnut and tara**,** respectively, and discussed in further detail below.

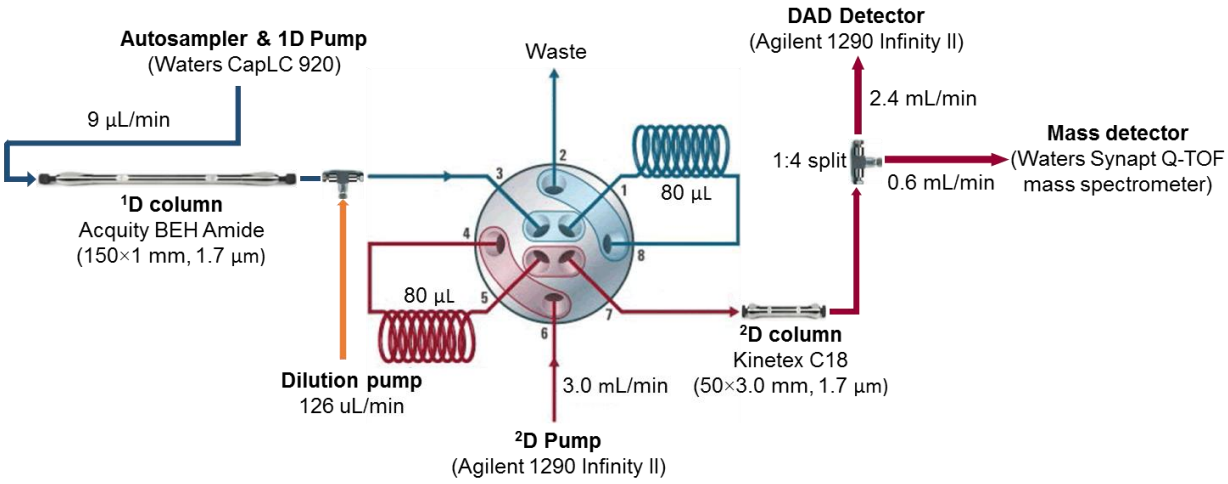


Figure 5.1: Schematic illustration of the instrumental set-up and conditions used for the comprehensive two-dimensional HILIC×RP-LC-UV-MS analysis for hydrolysable tannins. Figure adapted from [8].

5.3.2 On-line HILIC×RP-LC-MS analysis of chestnut and tara hydrolysable tannins

Identification of hydrolysable tannins in each of the analysed samples was based on HILIC and RP-LC retention times, UV and HR-MS data as well as IM arrival times compared to Chapters 3 and 4. The hydrolysable tannins identified are listed in **Table 5.2** for chestnut and **Table 5.3** for tara using the same compound numbers as in Chapters 3 and 4, respectively. A detailed discussion on the identification of individual compounds will not be presented in the present work; for a complete discussion on the tentative identification of the listed compounds in chestnut and tara, the reader is referred to Chapters 3 and 4.

The two-dimensional UV contour plots obtained for chestnut and tara are presented in **Figures 5.2** and **5.5**, respectively, with the corresponding base peak ion (BPI) 2D contour plots shown in **Figures 5.3** and **5.6**, respectively. From these figures, good utilisation of the two-dimensional separation space is evident for both samples, as is relatively good separation of individual compounds. A further noteworthy observation is the group-type separations obtained for the hydrolysable tannin species in both samples. This feature may facilitate the tentative identification of compounds based

on the positions in the two-dimensional space and may furthermore prove useful for the comparison of hydrolysable tannin profiles between different samples.

For the chestnut sample, three retention regions were identified for gallotannins, ellagitannins (monomers and dimers) and dimeric species containing both gallotannin and ellagitannin moieties (**Figure 5.2**). Representative structures for compound classes eluting in each of these three regions are shown in **Figure 5.4**: a trigalloyl-glucose isomer (**30, A**), vescalagin (**17[a], B**) and a vescalagin/castalagin-trigalloyl-glucose dimer (**29, C**).

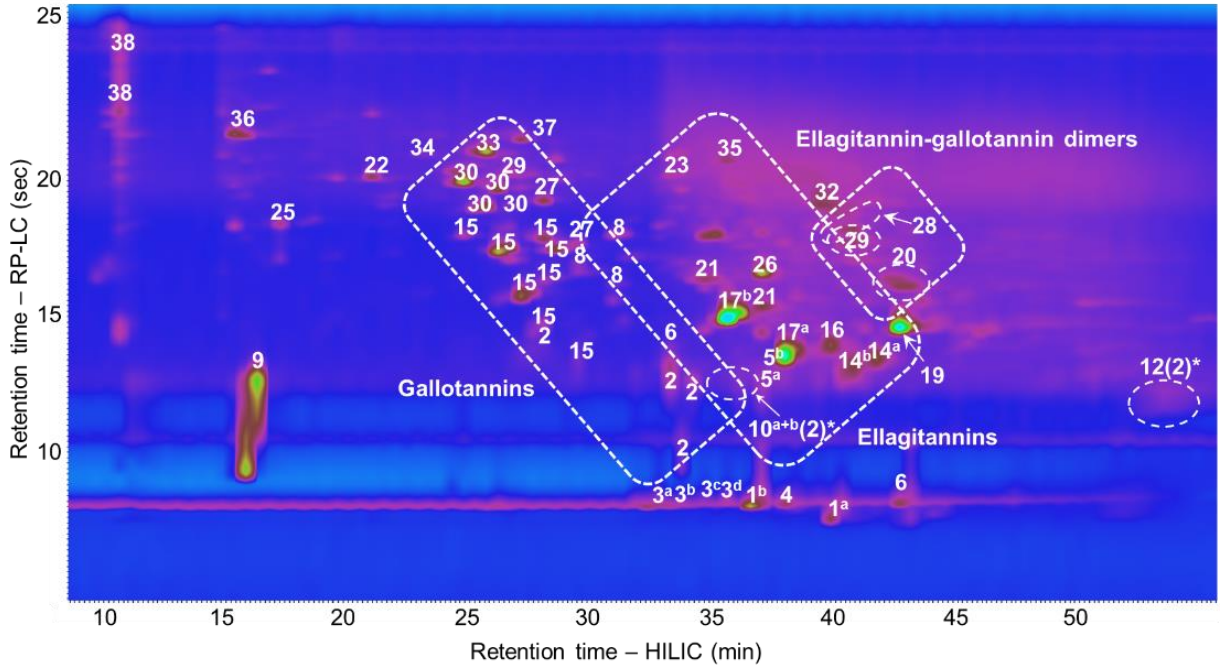


Figure 5.2: UV contour plot obtained at 280 nm for the on-line HILIC×RP-LC−MS analysis of a commercial chestnut sample. Peak numbers correspond to **Table 5.2**. Number in brackets ()* indicate the number of rotational or prototropic isomers separated by ion mobility spectrometry.

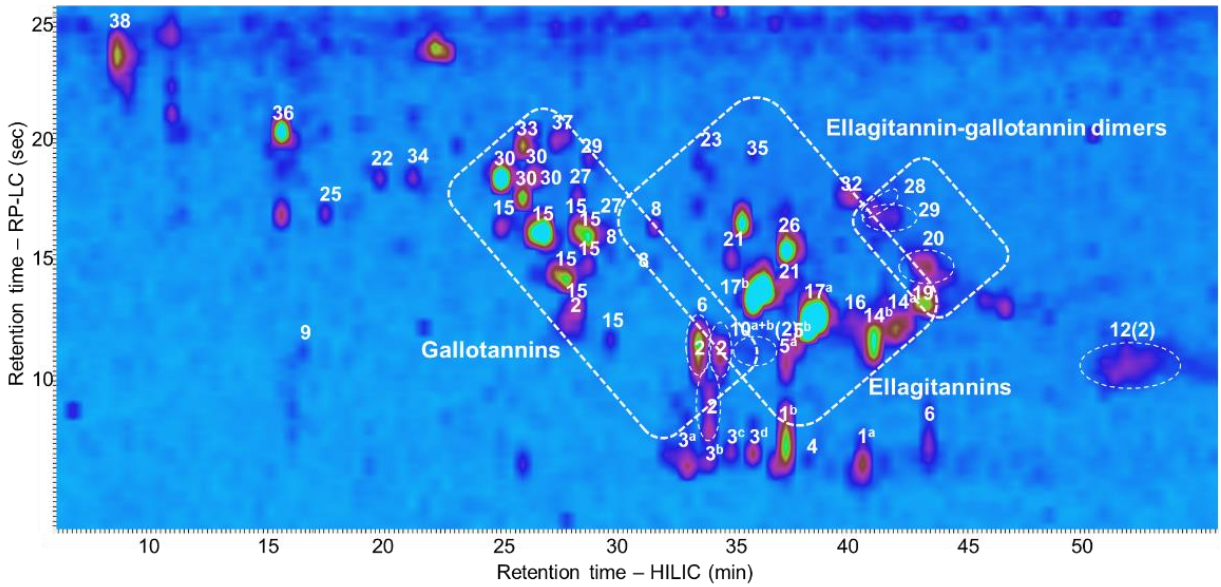


Figure 5.3: Base peak ion (BPI) contour plot obtained for the on-line HILIC×RP-LC−MS analysis of a commercial chestnut sample. Peak numbers correspond to **Table 5.2**. Numbers in brackets ()* indicate the number of conformers or prototropic isomers separated by ion mobility spectrometry.

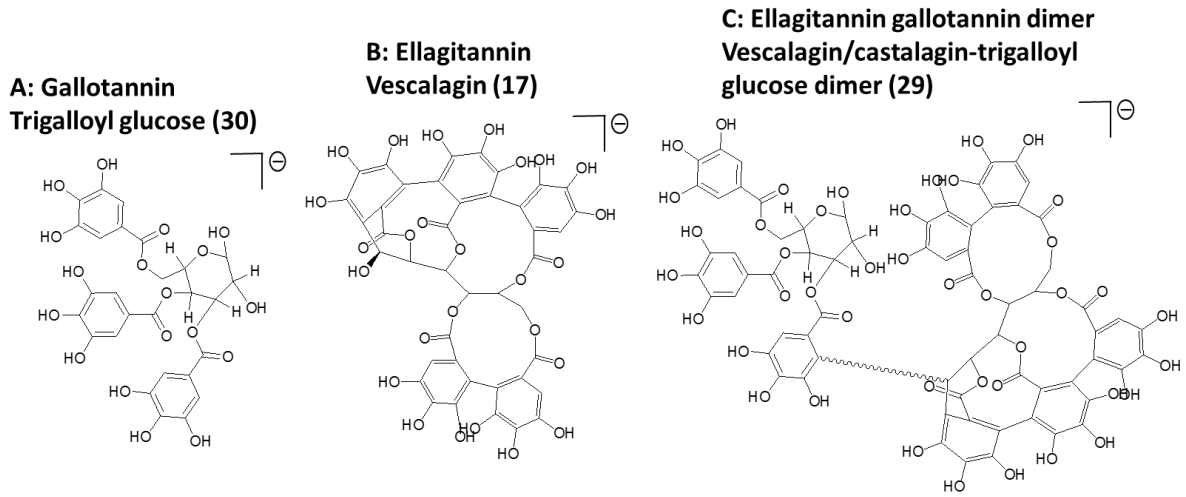


Figure 5.4: Representative structures of three compound classes distinguished by the group-type separation of chestnut hydrolysable tannins by HILIC×RP-LC-MS. Compound names with their respective numbers shown in brackets correspond to **Table 5.2**, and **A**, **B** and **C** refer to the relevant retention windows for each class indicated in **Figure 5.2**.

Table 5.2: Summary of the hydrolysable tannins detected in the chestnut sample by HILIC×RP-LC-MS analysis. Compound numbers correspond to Chapter 3 and **Figures 5.2-5.4**.

Compound (no)	Chemical formula	^{12}C Mass $[M-H]^{-1}$	Mass error (ppm)	t_R HILIC (min)	t_R RP-LC (sec)	UV (nm)	t_d*(ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)
Vescalin (1[a])	$C_{27}H_{20}O_{18}$	631.0585	2.15	39.95	7.68	231	3.80	220	nd*			
Castalin (1[b])	$C_{27}H_{20}O_{18}$	631.0582	1.67	36.66	7.99	232	3.80	220				
Monogalloyl glucose (2)	$C_{13}H_{16}O_{10}$	331.0664	-0.38	28.20	14.18	272	2.14	166	nd			
				33.37	12.99							
				33.84	10.04							
				34.31	12.56							
HHDP-Glucose (3[a-d])	$C_{20}H_{18}O_{14}$	481.0619	0.14	32.43	7.99	208	3.04	197	nd			
				33.84	7.95							
				34.78	8.11							
				35.72	8.11							
Pedunculagin - CO_2 (4)	$C_{33}H_{24}O_{20}$	739.0889	-7.85	38.07	8.15	204	4.90	250	nd			
Dihydrated Vescalagin (5[a])	$C_{41}H_{30}O_{28}$	969.0834	-1.14	37.60	12.51	230	6.35	284	484.0394	2.17	2.00	312
Di-hydrated Castalagin (5[b])	$C_{41}H_{30}O_{28}$	969.0790	-5.68	37.60	12.91	230	6.35	284	484.0393	1.96	2.00	312
Pedunculagin I (6)	$C_{34}H_{24}O_{22}$	783.0696	1.91	33.37	13.94	210	5.18	257	391.0322	3.84	1.66	284
				42.77	8.11							
Galloyl-HHDP-glucose (8)	$C_{27}H_{22}O_{18}$	633.0728	0.01	29.61	16.58	272	3.86	222	nd			
				31.02	16.18							
				31.49	17.80							
Gallic acid (9)	$C_7H_6O_5$	169.0128	-5.33	16.45	12.56	270	0.83	105	nd			
Roburin D (10[b])	$C_{82}H_{50}O_{51}$	nd		36.19	12.71	232	nd		924.0584	0.32	3.52	414
Roburin D (10[b])											4.35	464

Table 5.2 (continued).

Compound (no)	Chemical formula	^{12}C mass $[M-H]^{-1}$	Mass error (ppm)	t_R HILIC (min)	t_R RP-LC (sec)	UV (nm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)
C-galloylated vescalagin dimer (12)	$C_{89}H_{54}O_{55}$	nd		53.58	11.89	234	nd		1000.064	0.40	4.00	442
C-galloylated vescalagin dimer (12)											4.69	480
Grandinin (14[a])	$C_{46}H_{34}O_{30}$	1065.1068	1.06	41.83	13.29	nd	7.52	25.85	532.0522	6.20	2.21	328
Digalloyl glucose (15)	$C_{20}H_{20}O_{14}$	483.0765	-2.04	29.61	13.26	274	2.90	191	241.0343	-2.49	0.97	228
				28.20	14.69							
				27.26	15.71							
				28.20	16.18							
				28.67	17.44							
				28.20	17.88							
				26.32	17.36							
				24.91	17.95							
Vescavaloninic acid (16[a])	$C_{48}H_{30}O_{31}$	nd		39.95	13.81	224	8.00	319	550.0330	4.00	2.35	338
Vescalagin (17[a])	$C_{41}H_{26}O_{26}$	933.0604	-3.22	38.07	13.54	226	6.65	291	466.0284	1.29	1.93	306
Castalagin (17[b])	$C_{41}H_{26}O_{26}$	933.0633	-0.11	35.72	15.00	224	6.62	290	466.0293	3.22	1.93	306
1-O-galloyated castalagin (19)	$C_{48}H_{30}O_{30}$	1085.0804	5.55	42.77	14.53	220	7.80	315	542.0356	4.24	2.35	338
Vescalagin/castalagin-digalloyl glucose dimer (20)	$C_{61}H_{44}O_{39}$	nd		43.24	16.14	228, 280	nd		699.0663	1.57	2.97	381
Vescalene (21)	$C_{27}H_{18}O_{17}$	613.0466	0.03	34.78	16.26	252, 372	3.66	216	nd			
		613.0480	2.32	37.13	15.39	250, 372						
Valoneic acid dilactone (22)	$C_{21}H_{10}O_{13}$	469.0034	-1.92	21.15	20.16	252, 373	2.90	193	nd			

Table 5.2 (continued).

Compound (no)	Chemical formula	^{12}C mass $[M-H]^{-1}$	Mass error (ppm)	t_R HILIC (min)	t_R RP-LC (sec)	UV (nm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)
Punicalagin (23)	$C_{48}H_{28}O_{30}$	nd		33.37	20.11	238			541.0269	2.59	2.35	338
Digallic acid (25)	$C_{14}H_{10}O_9$	321.0231	-4.98	17.39	18.39	274	2.00	160				
1-O-trihydroxybenzene castalagin (26)	$C_{47}H_{30}O_{28}$	1041.0835	-1.54	27.13	16.43	228	7.38	306	520.0383	-1.15	2.28	333
Tellimagrandin I (27)	$C_{34}H_{26}O_{22}$	785.0801	-4.59	28.20 29.61	19.14 17.56	228, 274	5.45	263	392.0400	15.81	1.66	286
Vescalagin/castalagin-tetragalloyl glucose dimer (28)	$C_{75}H_{52}O_{47}$	nd		40.89	18.15	234, 280	nd		851.0766	0.59	3.45	410
Vescalagin/castalagin-trigalloyl glucose dimer (29)	$C_{68}H_{48}O_{43}$	nd		40.89	17.44	236, 280	10.42	364	775.0684	-2.97	3.17	393
Trigalloyl glucose (30)	$C_{27}H_{24}O_{18}$	635.0872	-1.96	25.85 26.79 26.32 24.91	18.98 19.10 19.65 20.00	276	3.86	220	317.0406	0.95	1.38	261
Vescalagin/castalagin-ellagic acid dimer (32)	$C_{55}H_{30}O_{33}$	nd		39.48	19.21	244, 372	8.35	325	608.0278	3.45	2.62	357
Tetragalloyl glucose (33)	$C_{34}H_{28}O_{22}$	787.0941	-6.73	25.85	21.10	278	5.52	264	393.0450	-2.04	1.73	297
Ellagic acid-pentose (34)	$C_{19}H_{14}O_{12}$	433.0413	1.39	23.50	21.03	nd	2.83	191				
Tellimagrandin II (35)	$C_{41}H_{30}O_{26}$	937.0977	3.20	35.72	20.63	236, 272	6.69	292	468.0455	4.27	2.00	312
Ellagic acid (36)	$C_{14}H_6O_8$	300.9994	3.32	15.51	21.61	254, 366	1.66	147	nd			
Pentagalloyl glucose (37)	$C_{41}H_{32}O_{26}$	939.1078	-2.77	27.26	21.50	278	6.76	0	469.0506	-1.49	2.14	325

Table 5.2 (continued).

Compound (no)	Chemical formula	^{12}C mass $[M-H]^{1-}$	Mass error (ppm)	t_R HILIC (min)	t_R RP-LC (sec)	UV (nm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)	Exp. m/z $[M-2H]^{2-}$	Mass error (ppm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)
Trimethyl ellagic acid (38)	$C_{17}H_{12}O_8$	343.0473	5.54	11.75	24.80	254, 366	2.21	170	nd			

* t_d: arrival time; nd: not detected.

For tara gallotannins, group-type separation was also observed (**Figures 5.5** and **5.6**). In this case, compounds are grouped according to their degree of galloylation (corresponding to the numbers in **Figures 5.5** and **5.6**, **Table 5.3**), with an increase in the number of gallic acid moieties generally resulting in a corresponding increase in both HILIC and RP-LC retention. Furthermore, within each group comprising regioisomers with the same degree of galloylation, partial differentiation is obtained according to the number of depsidic bonds (indicated by the numbers in brackets in **Figures 5.5** and **5.6**). Generally, an increase in the number of depsidic bonds results in lower HILIC and higher RP-LC retention, although this does not apply in all instances.

Also noteworthy for the tara sample is the detection of hexa- and hepta-galloyl quinic acid species (**6** and **7**, **Table 5.3**) using the HILIC×RP-LC method. These compounds were not reported in Chapter 4 based on 1D HILIC and RP-LC results. However, subsequent re-investigation of the 1D data confirmed the presence of these compounds. Further analysis of the hexa- and hepta-galloyl quinic acid species revealed the expected shoulder at 300 nm in their UV spectra (data not shown). Differentiation between isomeric species for hexa- and hepta-galloyl quinic acids based on MS^E data, or the extent of the absorbance shoulder at 300 nm, was inconclusive due to low abundance and poor chromatographic resolution of these species. Nevertheless, ion mobility data did reveal a slight increase in arrival times for earlier (HILIC) and later (RP-LC) eluting hexagalloyl quinic acid species, suggesting the presence of additional depsidic bonds for these compounds.

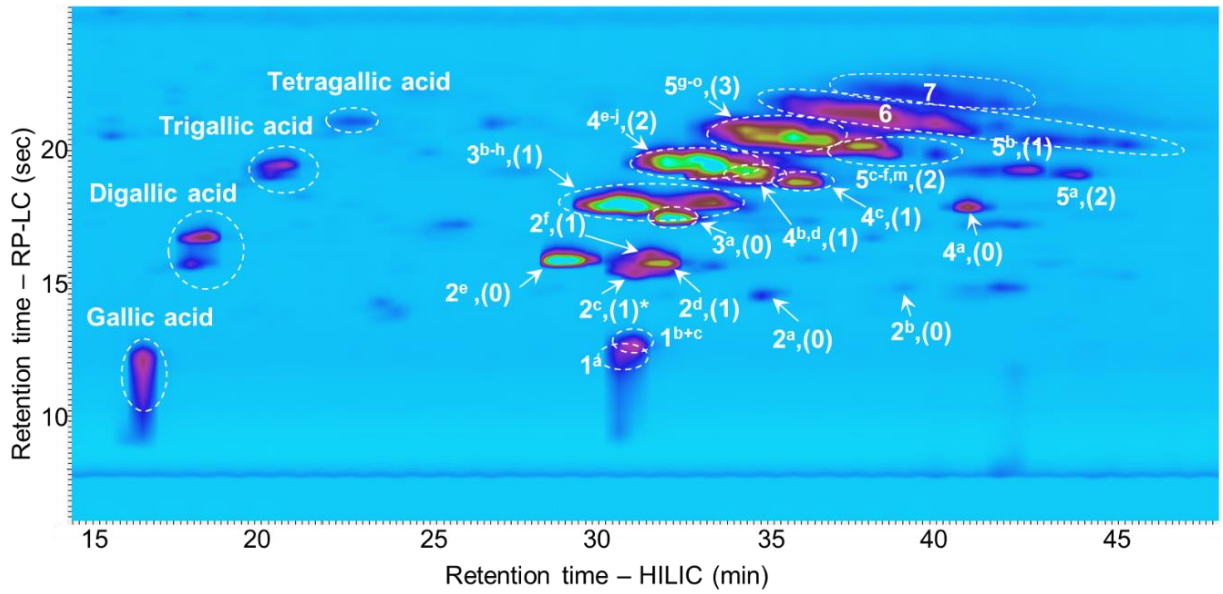


Figure 5.5: UV contour plot obtained for the HILIC×RP-LC−MS analysis of a commercial tara sample. Peak numbers correspond to **Table 5.3**. *Numbers in brackets indicate the number of depsidic bonds.

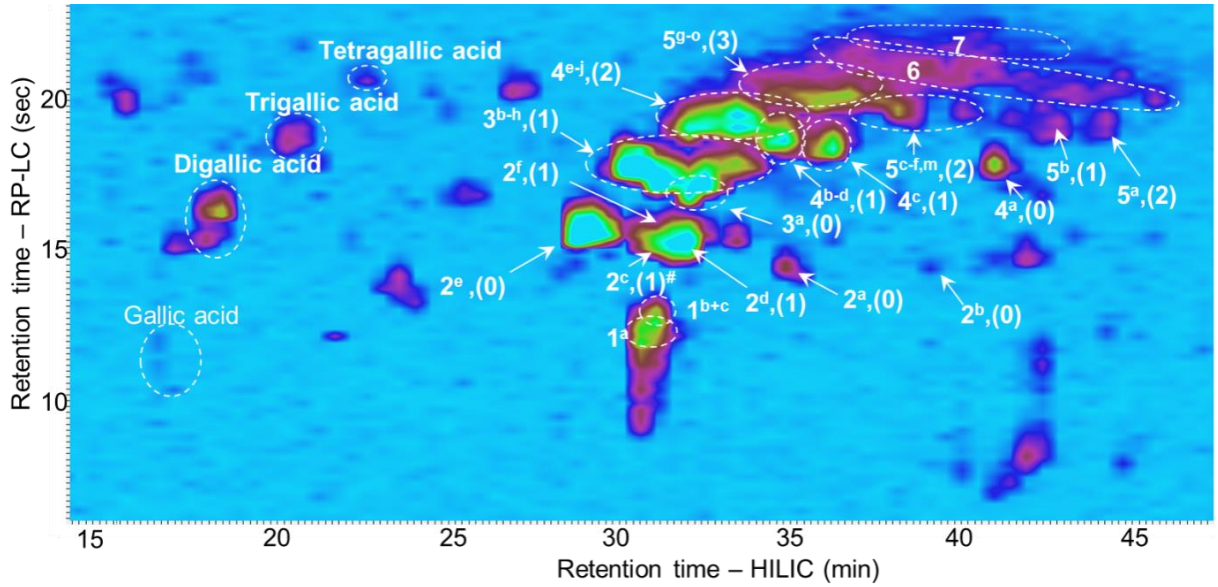


Figure 5.6: BPI contour plot obtained for the HILIC×RP-LC−MS analysis of a commercial tara sample. Peak numbers correspond to **Table 5.3**. *Numbers in brackets indicate the number of depsidic bonds.

Table 5.3: Galloyl-quinic acids and gallic acids tentatively identified in tara by HILIC×RP-LC-MS analysis. Compound numbers correspond to Chapter 4 and **Figures 5.5-5.6**.

Compound (no)	Chemical formula	^{12}C mass $[M-H]^{1-}$	Mass error (ppm)	t_R HILIC (min)	t_R RP-LC (sec)	UV (nm)	t_d* (ms)	$^{TW}CCS_{N2}$ (Å^2)
3-O-galloyl quinic acid (1[a])	$C_{14}H_{16}O_{10}$	343.0674	2.62	30.55	12.188	274	2.21	170
5 and 4-O-galloyl quinic acid (1[b+c])	$C_{14}H_{16}O_{10}$	343.0670	1.46	31.02	12.887	274	2.21	170
3,4-di-O-galloyl quinic acid (2[a])	$C_{21}H_{20}O_{14}$	495.0784	1.82	34.78	14.70	274	3.40	209
3,5-di-O-digalloyl quinic acid (2[b])	$C_{21}H_{20}O_{14}$	495.0777	0.40	39.01	15.05	274	3.40	209
3-O-digalloyl quinic acid (2[c])	$C_{21}H_{20}O_{14}$	495.0775	0.00	31.02	15.53	276[sh]*	3.10	199
							3.60	216
5-O-digalloyl quinic acid (2[d])	$C_{21}H_{20}O_{14}$	495.0794	3.84	31.96	15.93	274[sh]	3.10	199
							3.60	216
4,5-di-O-galloyl quinic acid (2[e])	$C_{21}H_{20}O_{14}$	495.0792	3.43	28.67	16.04	276	3.15	201
4-O-digalloyl quinic acid (2[f])	$C_{21}H_{20}O_{14}$	495.0776	0.20	31.49	16.24	274[sh]	3.30	206
							3.70	219
3,4,5-tri-O-galloyl quinic acid (3[a])	$C_{28}H_{24}O_{18}$	647.0887	0.46	31.96	17.61	276	4.00	226
Trigalloyl quinic acid (one depsidic bond) (3[b-h])	$C_{28}H_{24}O_{18}$	647.0903	2.94	30.08-33.84	18.13	273[sh]	4.35	237
		647.0904	3.09				4.35	237
		647.0908	3.71				4.35	237
		647.0890	0.93				4.35	237
1,3,4,5-tetra-O-galloylquinic acid (4[a])	$C_{35}H_{28}O_{22}$	799.0992	-0.25	40.89	18	274	4.76	247

Table 5.3 (continued).

Compound (no)	Chemical formula	^{12}C mass $[M-H]^{1-}$	Mass error (ppm)	t_R HILIC (min)	t_R RP-LC (sec)	UV (nm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)
Tetragalloyl quinic acid with one depsidic bond (4^{b+d})	$C_{35}H_{28}O_{22}$	799.1028	4.25	34.78	19.225	276^{sh}	5.04	253
Tetragalloyl quinic acid with one depsidic bond (4^{c})		799.1012	2.25	36.19	18.913		5.11	255
Tetragalloyl quinic acid (two depsidic bonds) (4^{e-j})	$C_{35}H_{28}O_{23}$	799.0996	0.25	31.49-33.84	19.7	274^{sh}	5.66	268
Pentagalloyl quinic acid (two depsidic bonds) (5^{a})	$C_{42}H_{32}O_{26}$	951.1125	2.21	44.18	19.225	274^{sh}	6.28	282
1-O-digalloyl-3,4,5-tri-O-galloylquinic acid (5^{b})	$C_{42}H_{32}O_{26}$	951.1116	1.26	42.77	19.425	274^{sh}	5.80	271
Pentagalloyl quinic acid (two depsidic bonds) ($5^{c-f,m}$)	$C_{42}H_{32}O_{26}$	951.1107	0.32	37.13-39.95	20.21	274^{sh}	6.28	282
Pentagalloyl quinic acid (three depsidic bonds) (5^{g-o})	$C_{42}H_{32}O_{26}$	951.1104	0.00	33.84-36.66	20.637	274^{sh}	7.04	299
Hexagalloyl quinic acid (6)	$C_{49}H_{36}O_{30}$	1103.1208	-0.45	36.19-46.06	20.32-21.97	274^{sh}	7.45	307
Heptagalloyl quinic acid (7)	$C_{56}H_{40}O_{34}$	1255.1316	-0.56	38.07-41.83	22.01-22.87	272^{sh}	8.49	328

Table 5.3 (continued).

Compound (no)	Chemical formula	^{12}C mass $[M-H]^{1-}$	Mass error (ppm)	t_R HILIC (min)	t_R RP-LC (sec)	UV (nm)	t_d (ms)	$^{TW}CCS_{N2}$ (Å^2)
Gallic acid (i)	$C_7H_6O_5$	169.0119	-13.61	1.72	271	co-elution	0.83	107
Digallic acid (ii)	$C_{14}H_{10}O_9$	321.0238	-4.36	16.45	12.16	270	2.00	162
				17.86	15.738			
Trigallic acid (iii)	$C_{21}H_{14}O_{13}$	473.036	-0.42	20.21	19.012	272^{sh}	3.17	202
				20.68	19.45	270^{sh}	2.9	193
Tetragallic acid (IV)	$C_{28}H_{18}O_{17}$	625.0478	1.12	22.56	21.075	270^{sh}	4.21	232
Ellagic acid (V)	$C_{14}H_6O_8$	300.9991	0.33	15.51	20.488	254, 366	1.66	147

* t_d: arrival time; sh: shoulder.

On the instrument used in this work, ion mobility (IM) data were also acquired in the HILIC×RP-LC-MS instrumental configuration used. This aspect will be addressed in detail in Chapter 6; in the present work IM data were primarily used as an additional identification criterion compared to data reported in Chapters 3 and 4. Nevertheless, some comments on this are warranted in the context of the LC×LC separations reported here.

Ion mobility coupled to the LC×LC workflow added no additional information not already obtained by 1D HILIC or RP-LC separations for the chestnut sample. For tara tannins, IM proved useful in the differentiation of positional isomers (as was the case in 1D LC, Chapter 4). Due to the unique arrival times of some isomers, these can be differentiated based on IM data alone, which, therefore, provides a fast way to assess their presence. This is illustrated in **Figure 5.7A**, which shows a combined mobilogram of the entire HILIC×RP-LC-MS analysis of the chestnut sample. The 3,4,5-tri-O-galloyl quinic acid (**3ª**) isomer detected at a unique arrival time of 4.00 ms, is easily assigned in this mobilogram, even allowing obtainment of clean mass spectra (**Figure 5.7B**).

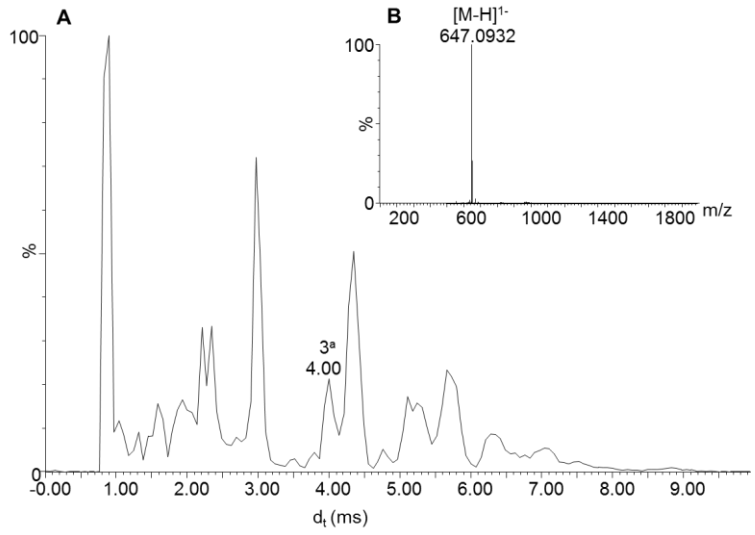


Figure 5.7: (**A**) Combined total ion mobilogram of the HILIC×RP-LC-MS analysis of the chestnut sample illustrating the unique arrival time for 3,4,5-tri-O-galloyl quinic acid (**3ª**) at 4.00ms. (**B**) shows the arrival-time mass spectrum of 3,4,5-tri-O-galloyl quinic acid.

5.3.3 Performance of on-line HILIC×RP-LC-MS for chestnut and tara hydrolysable tannins

As alluded to above, very fast ^{2}D analyses times (28 seconds in this work) are essential to minimise ^{1}D under-sampling in on-line LC×LC. This negatively impacts on resolution in the second dimension. Furthermore, on-column dilution in multidimensional chromatography can severely impact the sensitivity [11] of such separations. Despite the experimental approach used here, where the ^{1}D effluent is diluted using an aqueous make-up flow and large volumes are injected on the ^{2}D column, the high flow rate in the second dimension still results in significant dilution. These two factors combined limit the ultimate performance of on-line HILIC×RP-LC analysis for hydrolysable tannins. This is reflected in the number of isomeric species detected and/or resolved compared to 1D HILIC and RP-LC analyses. For example, in chestnut 128 isomeric species were detected by 1D RP-LC-MS (Chapter 3) compared to only 56 by HILIC×RP-LC-MS. Notably, roburins B and C (**11^{a}** and **11^{b}**, respectively) were not detected using the 2D LC method, whereas only one peak was detected for vescalagin/castalagin-digalloyl glucose (**20**) isomers compared to 16 observed in 1D RP-LC. This implies that further fine-tuning of the methodology might be required, specifically in terms of slightly longer modulation periods to benefit from improved ^{2}D RP-LC performance, at the cost of higher ^{1}D under-sampling.

In order to obtain an accurate estimate of the HILIC×RP-LC method performance, both under-sampling and orthogonality must be accounted for in the determination of the practical peak capacity ($n'_{c,2D}$). Briefly, under-sampling refers to the recombination of ^{1}D peaks due to a sampling times that are too long in relation to ^{1}D peak widths [12]. Orthogonality refers to the correlation of retention mechanisms between the first and second dimension separation modes, which determines the effective utilisation of the two-dimensional separation space [13]. Therefore, a weak correlation between the separation mechanisms results in good orthogonality, and *vice versa*.

To calculate the practical peak capacities of the separations reported here, ^{1}D peak capacities were estimated from the number of slices detected for selected compounds and the relevant modulation period (0.47 min). This provided practical ^{1}D peak capacities, corrected for under-sampling ($^{1}n'_{c}$), of 58 for chestnut and 65 for tara (**Table 5.4**). ^{2}D peak capacities were determined using the average peak widths and the

relevant gradient times for the same compounds in RP-LC. Peak capacity was corrected for limited surface coverage using the minimum convex hull method [14].

Using this approach, practical peak capacities, corrected for both under-sampling and orthogonality, of 1041 and 1263 for the two samples analysed here were obtained (**Table 5.4**). This confirms, to our knowledge for the first time, the potential of HILIC×RP-LC for the analysis of hydrolysable tannins.

Table 5.4: Performance parameters for the on-line HILIC×RP-LC-MS methodology for the separation of chestnut and tara hydrolysable tannins.

Performance parameter	Chestnut	Tara
HILIC practical peak capacity, $^{1}n'_{c}$	58	65
RP-LC peak capacity, $^{2}n_{c}$	29	29
[a]HILIC×RP-LC practical peak capacity, $n'_{c,2D}$	1707	1885
[b]Fractional surface coverage, $f_{c,2D}$ (%)	61	67
[c]HILIC×RP-LC practical peak capacity, $n''_{c,2D}$	1041	1263

[a]taking undersampling into account [12].
[b]calculated according to [14].
[b]taking undersampling [12] and orthogonality [13] into account

5.4 Conclusions

On-line HILIC×RP-LC methods were successfully developed and implemented in combination with UV and IM-MS detection for the comprehensive analysis of hydrolysable tannins, including ellagi- and gallotannins in chestnut and tara samples. Optimal HILIC×RP-LC conditions were derived by a predictive kinetic optimisation method, which included the use of an aqueous make-up flow to avoid injection band broadening in the second dimension. This approach provided practical peak capacities of more than a thousand (taking both undersampling and orthogonality into account). The structured elution orders observed for hydrolysable tannins in both samples proved especially beneficial in terms of tentative compound identification and demonstrate the promise of this methodology as a quick screening technique for hydrolysable tannins in natural products. Nevertheless, the inherent limitations of LC×LC in terms of dilution and ^{2}D resolution resulted in fewer compounds being detected in the analysed samples than was the case for 1D HILIC and RP-LC

analyses. Incorporation of ion mobility into the LC×LC workflow provided an additional identification criterion in terms of arrival time and allows the option of filtering mass spectra according to arrival time, but offered little additional separation information for hydrolysable tannins.

5.5 References

1. White T (1957) Tannins-Their occurrence and significance. J Sci Food Agric 8:377–385. doi: 10.1002/jsfa.2740080702

2. Okuda T, Yoshida T, Hatano T (1990) Oligomeric hydrolysable tannins, a new class of plant polyphenols. Heterocycles 30:1195–1218

3. Herve du Penhoat CLM, Michon VMF, Ohassan A, Peng S, Scalbert A, Gage D (1991) Roburin A, A dimeric ellagitannin from heartwood of Quercus robur. Phytochemistry 30:329–332. doi: 10.1016/0031-9422(91)84148-L

4. Cifuentes A, Bernal J, Herrero M, Ibá E (2009) Multidimensional chromatography in food analysis. J Chromatogr A 1216:7110–7129. doi: 10.1016/j.chroma.2009.08.014

5. Villiers A de, Venter P, Pasch H (2016) Recent advances and trends in the liquid-chromatography – mass spectrometry analysis of flavonoids. J Chromatogr A 1430:16–78. doi: 10.1016/j.chroma.2015.11.077

6. de Villiers A, Kalili KM (2017) Comprehensive Two-Dimensional Hydrophilic Interaction Chromatography × Reversed Phase Liquid Chromatography (HILIC×RP–LC): Theory, Practice, and Applications. In: Grushka E, Grinberg N (eds) Advances in Chromatography, Vol. 53. CRC Press,Taylor & Francis Group, pp 217–299

7. Willemse CM, Stander MA, Vestner J, Tredoux AGJ, de Villiers A (2015) Comprehensive Two-Dimensional Hydrophilic Interaction Chromatography (HILIC) × Reversed-Phase Liquid Chromatography Coupled to High-Resolution Mass Spectrometry (RP-LC-UV-MS) Analysis of Anthocyanins and Derived Pigments in Red Wine. Anal Chem 87:12006–12015. doi: 10.1021/acs.analchem.5b03615

8. Muller M, Tredoux AGJ, de Villiers A (2018) Application of Kinetically Optimised Online HILIC × RP-LC Methods Hyphenated to High Resolution MS for the Analysis of Natural Phenolics. Chromatographia. 82: 181-196.

9. Potts LW, Stoll DR, Li X, Carr PW (2010) The impact of sampling time on peak capacity and analysis speed in on-line comprehensive two-dimensional liquid chromatography. J Chromatogr A 1217:5700–5709. doi: 10.1016/j.chroma.2010.07.009

10. Muller M, Tredoux AGJ, de Villiers A (2018) Predictive kinetic optimisation of hydrophilic interaction chromatography × reversed phase liquid chromatography separations: Experimental verification and application to phenolic analysis. J Chromatogr A 1571:107–120. doi: 10.1016/j.chroma.2018.08.004

11. Schure MR (1999) Limit of detection, dilution factors, and technique compatibility in multidimensional separations utilizing chromatography, capillary electrophoresis, and field-flow fractionation. Anal Chem 71:1645–1657. doi: 10.1021/ac981128q

12. Murphy RE, Schure MR, Foley JP (1998) Effect of Sampling Rate on Resolution in Comprehensive Two-Dimensional Liquid Chromatography. Anal Chem 70:1585–1594. doi: 10.1021/ac971184b

13. Schure MR, Davis JM (2017) Orthogonality measurements for multidimensional chromatography in three and higher dimensional separations. J Chromatogr A 1523:148–161. doi: 10.1016/j.chroma.2017.06.036

14. Semard G, Peulon-Agasse V, Bruchet A, Bouillon JP, Cardinaël P (2010) Convex hull: A new method to determine the separation space used and to optimize operating conditions for comprehensive two-dimensional gas chromatography. J Chromatogr A 1217:5449–5454. doi: 10.1016/j.chroma.2010.06.048

Declaration with signatures in possession of candidate and supervisor.

Declaration by the candidate:

With regard to Chapter 6, the nature and scope of my contribution were as follows:

Nature of contribution	**Extent of contribution (%)**
Performed the experiments, data analysis, co-wrote paper	40

The following co-authors have contributed to Chapter 6:

Name	**E-mail address**	**Nature of contribution**	**Extent of contribution (%)**
Magriet Muller	mullerm@sun.ac.za	Experimental set-up, data analysis; editorial input	10
Jochen Vestner	jochen.vestner@gmx.de	Developed the data visualization protocol, prepared figures; editorial input	10
Maria A. Stander	lcms@sun.ac.za	Assisted with experimental set-up and data manipulation; editorial input	5
Andreas G.J. Tredoux	atredoux@sun.ac.za	Editorial input	5
Harald Pasch	hpasch@sun.ac.za	Editorial input	5
André de Villiers	ajdevill@sun.ac.za	Co-wrote paper	25

Signature of candidate:

Date: 25\1\2018

Declaration by co-authors:

The undersigned confirm that:

1. The declaration above accurately reflects the nature and extent of the contributions of the candidate and the co-authors to Chapter 6,
2. No other authors contributed to Chapter 6 besides those specified above, and
3. Potential conflicts of interest have been revealed to all interested parties and that the necessary changes have been made to use the material in Chapter 6 of this dissertation.

Signature	Institutional affiliation	Date
	Stellenbosch University	25/1/2018
	DLR Rheinpfalz	25/1/2018
	Stellenbosch University	25/1/2018
	Stellenbosch University	25/1/2018
	Stellenbosch University	25/1/2018
	Stellenbosch University	25/1/2018

Chapter 6

Comprehensive 3-dimensional LC×LC×ion mobility spectrometry separation combined with high resolution MS for the analysis of complex samples[#]

Abstract

Comprehensive two-dimensional liquid chromatography (LC×LC) and ion mobility spectrometry−mass spectrometry (IM−MS) are increasingly being used to address challenges associated with the analysis of highly complex samples. In this work, we evaluate the potential of the combination of these techniques in the form of a comprehensive three- dimensional LC×LC×IM separation system. As application, hydrophilic interaction chromatography (HILIC) × reversed phase LC (RP-LC)×IM high-resolution MS (HR-MS) was used to analyse a range of phenolic compounds, including hydrolysable and condensed tannins, flavonoids, and phenolic acids in several natural products. A protocol for the extraction and visualisation of the four-dimensional data obtained using this approach was developed. We show that the combination of HILIC, RP-LC, and IM offers excellent separation of complex phenolic samples in three dimensions. Benefits associated with the incorporation of IM include improved MS sensitivity and mass-spectral data quality. IM also provided separation of trimeric procyanidin isomeric species that could not be differentiated by HILIC×RP-LC or HR-MS. On the traveling wave IMS (TWIMS) system used here, both IM separation performance and the extent of second dimension (^{2}D) undersampling depend on the upper mass scan limit, which might present a limitation for the analysis of larger molecular ions. The performance of the LC×LC×IM system was characterised in terms of practical peak capacity and separation power, using established theory and taking undersampling and orthogonality into account. An average increase in separation performance by a factor of 13 was found for the samples analysed here when IM was incorporated into the HILIC×RP-LC−MS workflow.

[#]This chapter has been published in Anal. Chem. 2018, 90, 11643-11650

6.1 Introduction

The demand for high-resolution separations of highly complex mixtures, such as those encountered in proteomics, metabolomics, and natural product analysis, has provided the impetus for rapid development in multidimensional liquid chromatographic (MDLC) separations and comprehensive two-dimensional LC (LC×LC) in particular. On-line LC×LC offers significant performance gains at conventional analysis times, whereas widespread application of the technique has been facilitated by the availability of powerful commercial instrumentation.

The ultimate performance of on-line LC×LC is however constrained by the inherent limitations of LC for very fast second dimension (^{2}D) separations dictated by sampling requirements to maintain first dimension (^{1}D) separation performance. Currently, the maximum peak capacity attainable by LC×LC for reasonable analysis times is in the region of a few thousand. Considering that the available peak capacity should significantly exceed the number of compounds to be separated [1] and that the most complex samples of interest may contain in excess of tens of thousands of compounds (e.g., protein tryptic digests) [2,3], the need for further improvement in separation performance is clear. This is the context for renewed interest in comprehensive three-dimensional (3D) separations [4-7]. Moore and Jorgenson reported a 3D size-exclusion chromatography (SEC) × reversed phase LC (RP-LC) × capillary zone electrophoresis (CZE) separation of peptides in a landmark paper in 1995 [8]. However, this and further fundamental studies [4-6] also highlighted the ramifications of on-line time-based 3D-LC (^{t}LC×^{t}LC×^{t}LC) systems: sampling constraints dictate decreasing performance gains for each subsequent separation, accompanied by significant dilution. Combination of time-based and spatial (^{x}LC) separations appears to be a more realistic means of a further order of magnitude gain in peak capacity [5−7], although this has yet to be demonstrated in practice.

Of course, hyphenation of LC to MS somewhat relaxes the demands placed on chromatographic separation of species because of the inherent selectivity of MS detection, especially in the form of tandem or high resolution (HR) MS. However, MS also suffers from some limitations, such as distinguishing between isomeric species and its susceptibility to matrix effects.

The addition of ion mobility (IM) spectrometry to LC−MS workflows offers a means of improving their performance for complex sample analysis. IM entails the gas-phase separation of ions in an inert buffer gas in the presence of an electric field [9,10]. Benefits associated with IM−MS include improved spectral quality in data-independent acquisition (DIA) modes from the filtering of MS or MS/MS data according to arrival time [10], increased duty cycle [11], separation of ions of different charge states, additional fragmentation options [12], and the potential separation of isomeric and isobaric compounds [13]. Furthermore, experimentally determined collisional cross section (CCS) values can be used to obtain conformational and gas-phase interaction information [14] and may be used as an additional parameter for compound identification [15].

Although the resolving power of current IM instruments is insufficient for the separation of complex mixtures [16], the fact that these separations are attained on a millisecond time scale makes the technique promising in the context of comprehensive MD separations. The fast cycle times of IM separations allows the technique to meet the sampling requirements of even very fast ^{2}D separations, such as those encountered in on-line LC×LC, while at the same time providing acceptable separation performance (roughly similar to the CZE separation used by Moore and Jorgenson in the third dimension) [8]. Indeed, the utility of IM-MS combined with MDLC separations has been demonstrated by Schmitz and co-workers [17,18]. These authors used drift tube (DT) IM−HR-MS combined with “continuous multiple heart-cutting” LC separation, in which long modulation periods were used to facilitate data interpretation using commercial software. More recently, the combination of traveling wave IM (TWIMS) with comprehensive hydrophobic interaction chromatography (HIC) × SEC separation for the characterisation of antibody-drug conjugates was reported [19], although the ^{2}D SEC separation was primarily used as a desalting step.

In this work we explore in detail the performance and potential of comprehensive 3D LC×LC×IM [19−21] hyphenated with high-resolution MS. Regarding the nomenclature used in this work, we refer to the combination of LC×LC with IM as a comprehensive 3D LC×LC×IM separation even though the compounds separated by LC×LC are ionised prior to their separation by IM. This is in accordance with established nomenclature for comprehensive separations when soft ionisation, such as ESI, is

used [20]. Furthermore, although the combination of such a system with HR-MS meets the criteria for being considered a comprehensive four-dimensional separation [19,20], we treat the MS as a detector in the present work [i.e., not as a separate separation step]. As application, the combination of on-line comprehensive HILIC×RP-LC separation with TWIMS−HR-MS for the analysis of a variety of complex natural phenolic mixtures is used. The performance of the system is discussed critically in terms of the benefits offered by IM and the separation performance practically achievable considering instrumental and software constraints. In order to overcome one of the main challenges of such a system, namely, the lack of commercial software to accommodate the resulting four-dimensional data, a protocol for the visualisation and interpretation of such data using freely available software is also reported.

6.2 Experimental section

6.2.1 Reagents and materials

HPLC-grade acetonitrile (ACN), methanol (MeOH), and formic acid (FA) as well as poly-DL-alanine were purchased from Sigma-Aldrich (Johannesburg, South Africa). Deionied water was obtained using a Milli-Q water purification system (Millipore, Milford, MA). The commercial chestnut tannin extract was provided by Silvateam (San Michele Mondovì, Italy). The red wine sample (Vitis vinifera Pinotage varietal, 2016 vintage) was purchased commercially, and grape seeds (Vitis vinifera,Cabernet Sauvignon) of the 2011 harvest were obtained from the Institute of Wine Biotechnology (IWBT, Stellenbosch University, Matieland, South Africa). The rooibos tea sample was a fermented commercial sample of Aspalathus linearis [22].

6.2.2 Sample preparation

The chestnut extract (49 mg) was dissolved in 1 mL of MeOH/H_2O (50:50, v/v) and filtered through a 0.45 µm hydrophilic PVDF membrane (Millipore). Phenolic extracts were prepared as previously reported for red wine [23] grape seed [24] and rooibos tea [22] samples and dissolved in different ratios of MeOH/H_2O or MeOH/ACN as specified in **Table S-1** in the Supporting Information (SI).

6.2.3 Instrumentation

A Waters Capillary LC 920 pump and autosampler equipped with a 1 or 2 µL loop and controlled by MassLynx v4.1 software (Waters, Milford, MA) was used in the ^{1}D separation. The ^{2}D separation was performed on an Agilent 1290 Infinity binary pump connected via a two-position, eight-port valve equipped with two 80 µL loops and controlled by OpenLab CDS software (Agilent Technologies, Waldbronn, Germany). An Agilent 1100 Isocratic pump provided a makeup flow of pure water via a T-piece placed between the ^{1}D column and the valve. The ^{2}D column was thermostated to 50 °C in an Agilent 1200 column oven. After the ^{2}D column, the flow was split 4:1 between an Agilent 1290 Infinity II DAD detector (1 µL flow cell, 190−640 nm, 80 Hz) and a Waters Synapt G2 quadrupole time-of-flight (Q-TOF) mass spectrometer equipped with an ESI source.

6.2.4 Chromatographic conditions

^{1}D HILIC separations were performed using an Acquity UPLC BEH amide column (150 × 1 mm i.d., 1.7 µm, Waters) at ambient temperature, and ^{2}D RP-LC separations were performed on a Kinetex C_{18} core−shell column (50 × 3.0 mm i.d., 1.7 µm, Phenomenex, Torrance, CA). Mobile phases comprised (A) 0.1% (v/v) FA in water and (B) ACN in both dimensions. Details on the mobile-phase flow rates, gradients, ^{1}D injection volumes, dilution flows, and modulation periods used for each of the samples are summarised in **Table S-1**.

6.2.5 IM-MS conditions

MS data were acquired in negative ionisation mode with capillary, cone, and extraction-cone voltages of −2.5 kV, 15.0 V, and 4.0 V, respectively, and source and desolvation temperatures of 120 and 275 °C, respectively. Nitrogen was used as the desolvation gas at a flow rate of 650 L/h. Data were acquired using a scan time of 0.2 sec (5 Hz) from 100 to 2000 amu, except for the data from the sensitivity experiments (**Section S-1**), in which both the mass range and the scan time were varied as indicated. The instrument was calibrated using a sodium formate solution, and leucine enkephalin (m/z 554.2615, $[M-H]^{1-}$) was used as the lock mass calibrant.

For IM measurements, nitrogen was used as the drift gas at a flow of 90 mL/min (3.27 mbar), and a He flow of 180 mL/min (1410 mbar) was used for the He cell. The mobility

T-wave was operated at a constant velocity of 448 m/s and a wave height of 37.1 V. The transfer velocity was 380 m/s. The trap collision energy, trap DC bias, and helium-cell DC voltages were set to 7, 45, and 35 V, respectively.

To determine experimental CCS values (Ω) from the measured arrival times, calibration was performed using poly-DL-alanine as calibrant [25]. The accuracy of the calibration was confirmed by comparison of CCS values of several phenolic compounds to literature values and values obtained on a DTIMS instrument (**Table S-2**).

6.2.6 Data processing

For conventional data analysis and visualisation, commercial software was used: UV and MS contour plots were constructed using LC Image software (v2.6, GC Image LLC, Lincoln, NE). MS and IM data were processed in MassLynx v4.1 and Driftscope v2.1 software (Waters).

For visualisation and interpretation of four-dimensional LC×LC×IM-MS data, data files were converted from the vendor format (.raw) to mzML format using the MSConvert tool from the freely available ProteoWizard [26] software. All further data processing was done in Python [27] and IPython [28] using in-house written scripts and Jupyter Notebooks. The Python package Pyteomics [29] was used to import mzML files. Imported data were handled in Pandas DataFrames [30] and saved to disk as pickles. Visualization was realized using the Plotly's Python graphing library [31]. All data processing and calculations were done on a HP Z240 Tower Workstation with an Intel Core i7-6700 with 3.4 GHz and 64 GB of DDR4 ram.

6.3 Results and discussion

6.3.1 Separation and data presentation

In order to evaluate the performance of IM−MS combined with LC×LC, several natural product samples containing complex mixtures of phenolic compounds were analysed. For the chromatographic separation, on-line HILIC×RP-LC was used because of the well-documented performance of such systems for phenolics [32,33]. Predictive Pareto optimisation of HILIC×RP-LC separations was used to derive optimal

conditions, as is reported elsewhere [34]. Briefly, HILIC separations on an amide column were combined with RP-LC separation on a C_{18} 1.7 µm core−shell column. To avoid injection band broadening due to the transfer of the HILIC mobile phase to the ^{2}D stage, dilution of the ^{1}D effluent with a makeup flow of water was performed [35]. Experimental details are summarised in **Table S-1**, and a schematic representation of the instrumental setup is presented in **Figure S-1**.

As application, four samples containing a range of different phenolic compounds were analysed; these include hydrolysable (ellagitannins and gallotannins in chestnut) and condensed tannins (procyanidins in grape seed) [24] as well as a range of flavonoid and nonflavonoid phenolics in rooibos tea [22] and red wine [23]. Examples of the UV contour plots obtained for each of these samples are presented in **Figure 6.1**, where it is clear that good chromatographic performance in terms of peak widths, sensitivity, and surface coverage is obtained for all samples. MS contour plots for the same samples are shown in **Figure S-2**. Further details on the separation and identification of compounds in each of these samples are reported elsewhere [36].

Combining the LC×LC separation with IM results in an additional separation step prior to MS detection, where sampling requirements for comprehensive separation are met because of the fast IM separation (see below). An example of the resulting data in the form of an (LC×)LC×IM heat map is shown in **Figure S-3**, with an insert illustrating the ^{2}D separations obtained during two consecutive modulation periods.

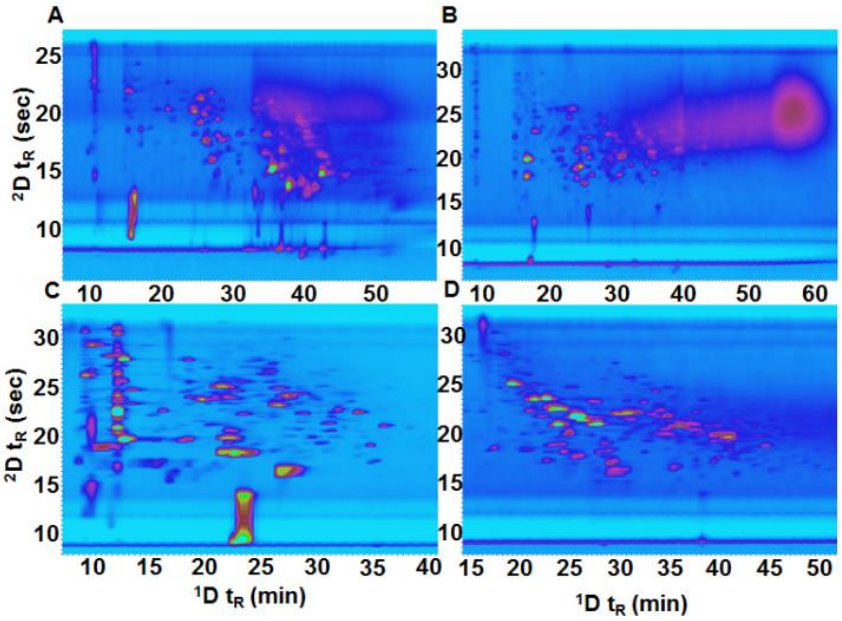


Figure 6.1: Examples of two-dimensional UV (280 nm) contour plots obtained for the on-line HILIC×RP-LC separation of (**A**) chestnut, (**B**) grape seed, (**C**) wine and (**D**) rooibos tea phenolics. For experimental conditions, refer to **Table S-1**.

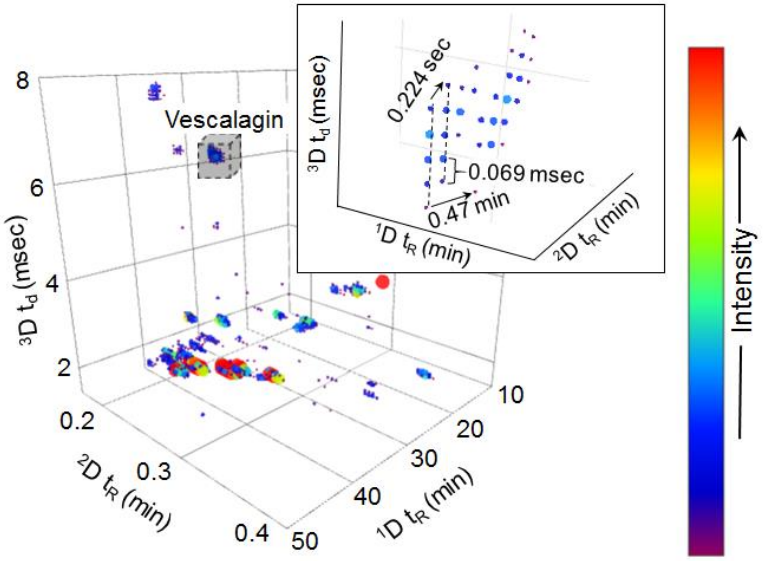


Figure 6.2: Three-dimensional representation of the HILIC×RP-LC×IM−MS TIC data obtained from the analysis of chestnut hydrolysable tannins. The inset shows an enlargement of the data points for the highlighted peak (vescalagin). For an interactive version of this figure, refer to the AVI video file uploaded as SI.

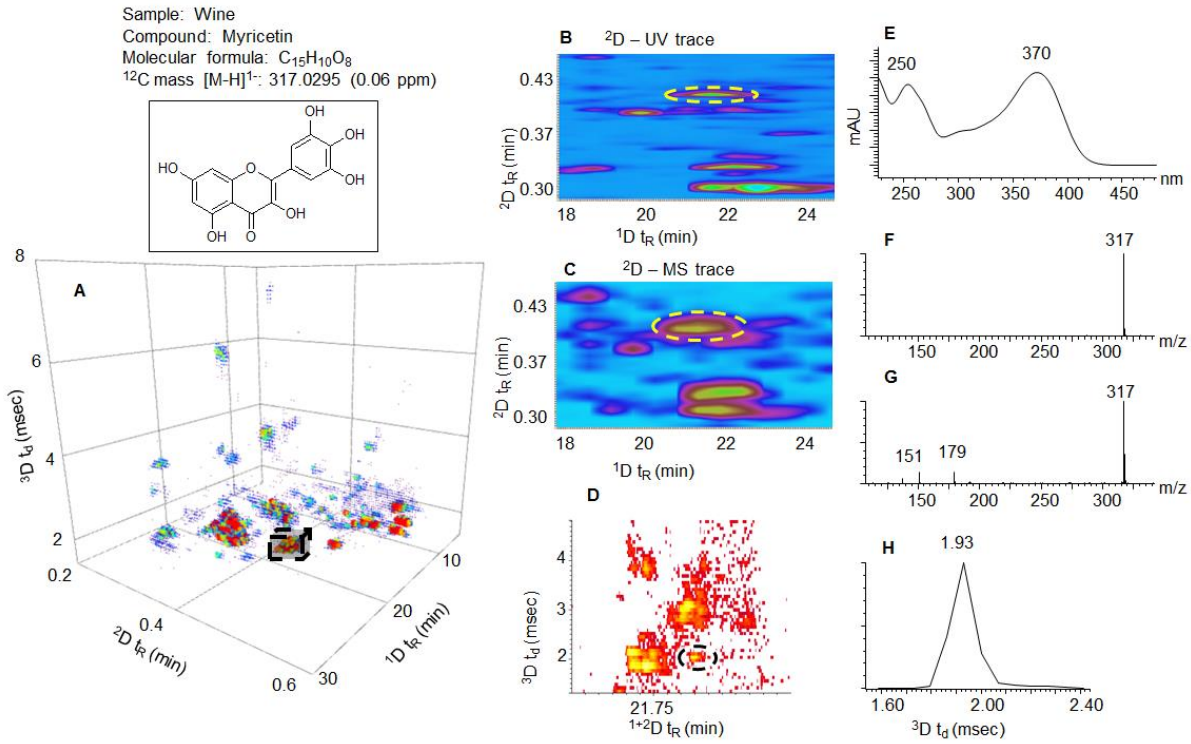


Figure 6.3: Example of the identification of the flavonol myricetin in a red wine sample based on HILIC×RP-LC-UV × IM−MS data. (**A**) Three-dimensional representation of the data. (**B**) UV, (**C**) TIC, and (**D**) drift-time contour plots of the relevant separation window. (**E**) UV spectrum, (**F**) low- and (**G**) high-collision energy mass spectra, and (**H**) extracted-ion mobilogram for myricetin. The position of myricetin is indicated by the dotted black square and the yellow and black circles. For further details on this compound, refer to **Table S-2** in the SI.

Commercial software, however, does not allow representation of the four-dimensional data in an easily interpretable manner. Therefore, in order to visualise the data generated by the HILIC×RP-LC×IM−MS system, a protocol using freely available software to manipulate exported raw data was developed (Experimental section).

This allowed us to generate an interactive display of the data, where data are plotted in three dimensions representing the ^{1}D HILIC, ^{2}D RP-LC, and ^{3}D IM separations, and peak intensity is indicated by a chosen colour scale. Mass-spectral data can be accessed for each data point using a cursor. An example of such an interactive display of the data can be found in the AVI video file uploaded as Supporting Information,

whereas **Figure 6.2** presents a static display of the data obtained for the chestnut sample.

The insert in **Figure 6.2** illustrates in more detail the format of the data for a particular compound, vescalagin. Each peak (i.e., three-dimensional spot) comprises multiple slices in the x (^1t_R), y (^2t_R), and z (^3t_d) directions, with each slice made up of multiple data points. Individual data points are separated by 0.47 min in the first dimension, which is equivalent to the modulation period used. Similarly, in the second dimension (RP-LC), data points are separated by 0.224 sec, corresponding to the MS cycle time (see further discussion on this below), and in the third dimension (IM), data points are 0.069 ms apart, corresponding to the TOF pusher frequency.

For the identification of compounds in each of the analysed samples, retention information in both the HILIC and RP-LC dimensions, on-line UV-spectral data, and low- and high- collision-energy HR-MS data can be used for conventional data analysis using commercial software. In addition, the measured arrival times (t_d) for each compound can be used for further confirmation of tentatively assigned compounds [15,37]. (Note that because commercial software is currently not able to accommodate LC×LC×IM−MS data, data analysis was performed manually in the present work; the labor-intensive nature of this process represents an obvious drawback). As an example, **Figure 6.3** illustrates the tentative identification of the flavonol aglycone myricetin in the red wine sample on the basis of the above-mentioned data. Similar examples of the identification of selected compounds in each of the other samples analysed are presented in **Figures S-4, S-5** and **S-6**, and the data for each of the examples are summarised in **Table S-2** [22,38−40]. One of the primary benefits of incorporating IM is that this allows MS data to be filtered according to arrival time, thereby providing cleaner mass spectra for partially coeluting compounds, as illustrated in **Figure S-7**. This figure shows how IM can partially compensate for the effect of inevitable extra-column band broadening on the MS signal (provided that the closely eluting species have different arrival times). In terms of separation, IM is only realistically expected to provide significant additional information for isomeric compounds, which cannot be distinguished by HR-MS, coelute following chromatographic separation, and can be separated on the basis of differences in their CCS values. A good example of the complementary separation

mechanism offered by IM is the procyanidin trimeric species (m/z 865, $[M-H]^-$) found in grape seed.

Oligomeric procyanidins are exceedingly hard to analyse because the number of isomeric species increases exponentially with the degree of polymerisation (DP) [24]. **Figure 6.4A** shows the data obtained for the HILIC×RP-LC-UV×IM−MS analysis of the grape seed sample, with the data for the procyanidin trimers highlighted in the insets (**Figure 6.4B−D**). From the relevant part of the extracted-ion contour plot in **Figure 6.4B**, four procyanidin trimers can be observed, which are relatively well-separated as a result of the combination of HILIC and RP-LC selectivities.

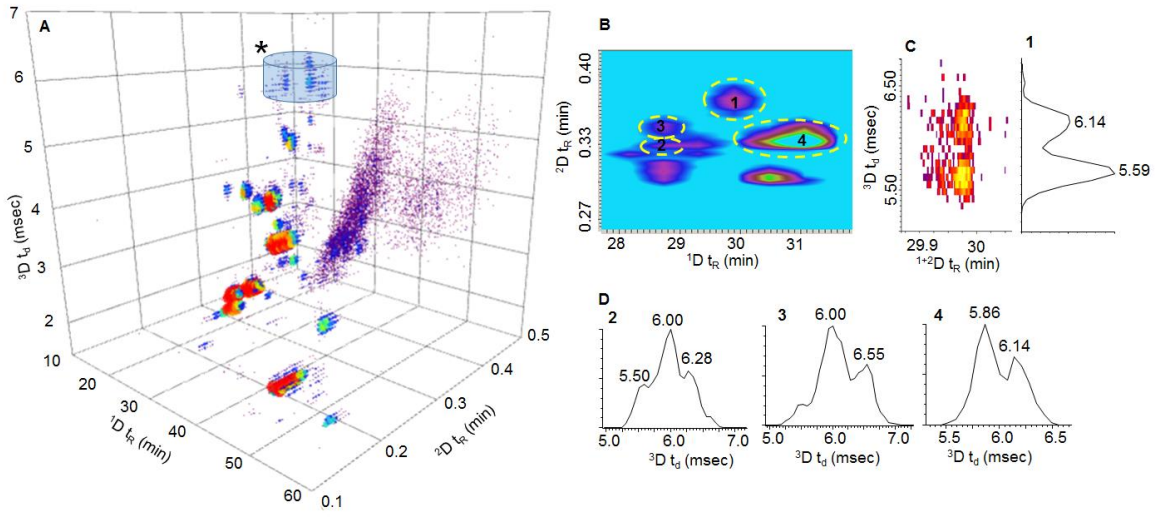


Figure 6.4: Three-dimensional separation (HILIC×RP-LC×IM-MS) of the procyanidin trimers (m/z 865) found in grape seed. **(A)** Shows the three-dimensional representation of the data, **(B)** the extracted ion contour plot for m/z 865 and **(C,D)** show the mobilograms obtained for the peaks labelled **1-4** on the extracted ion contour plot. *Indicates the three-dimensional separation space occupied by the procyanidin trimeric species.

However, the extracted-ion mobilograms for each of these peaks (**Figure 6.4 C,D**) show multiple peak maxima in each case, most evident for the peak labeled 1. Each of these sets of peaks observed in the drift dimension coeluted following HILIC×RP-LC separation and could not be distinguished on the basis of HR-MS data. In this

example, IM therefore provides additional information not accessible by LC LC−HR-MS analysis. At this point it is not certain what each of the sets of peaks separated by IM represent; we suspect they either correspond to rotational conformers [41]; prototropic ions [42] differing in terms of their deprotonation sites; or, less likely because these species are normally chromatographically separated, 4,6- and 4,8-linked isomers [43].

6.3.2 Evaluation of IM performance as a function of analytical conditions

Apart from the complementary separation capability gained by incorporation of IM into LC×LC−MS, both chromatographic and MS data are also affected in a manner that depends on the instrumental operating conditions. We therefore explored the influence of TWIMS−MS conditions on the performance of LC×LC×IM−MS systems in more detail by injecting a solution of poly-DL-alanine and acquiring MS data both with and without IM activated for different scan ranges and scan times. From the signal intensities (ion counts) and noise levels measured for each scan range, a substantial increase in the duty cycle of the instrument in IM−MS mode for upper mass limits up to about m/z 1000 was observed (**Figures S-8** and **S-9**). This is a consequence of synchronisation of the arrival time of ion packets from the IM cell, and the TOF pusher, which increases the duty cycle for a particular m/z range (for the settings used here, around m/z 585) [44]. For further details on the factors affecting sensitivity, the reader is referred to Section S-1 in the SI.

From the perspective of comprehensive 3D separations, the speed (cycle time, $t_{d,cycle}$) and peak widths obtainable by IM are particularly relevant. These can also be assessed using the data for poly-DL-alanine, obtained as outlined above, and from the manner in which the data are compiled by the software. In terms of the cycle time of TWIMS separations, these are inversely related to the upper mass of the scan range. For example, on the instrument used here (Synapt G2), the IM cycle time varies between 7.5 and 22 ms for upper mass ranges of ≤600 and >2000, respectively (**Table S-3**). This is because each mobilogram comprises 200 data points, or individual TOF push events, whereas the pusher frequency is inversely related to the highest mass of the scan range. Therefore, TWIMS separations are more than sufficiently fast to maintain the information provided by even the fastest ^{2}D separations in LC×LC.

However, on this instrument, all IM separations corresponding to the MS scan time are summed and saved by the software as a single mobilogram. This means that the effective IM cycle time corresponds to the scan time, which has important implications for LC×LC×IM−MS separations. Although TOF detectors are sufficiently fast to be used in LC×LC, sensitivity does drop off beyond a certain acquisition rate, depending on the scan range. On the instrument used here, for example, this occurs for scan times faster than 0.2 sec for an upper mass range of m/z 2000. Although this scan time will inevitably lead to some additional broadening of very narrow ^{2}D peaks (**Figures S-2** and **S-7**) and therefore potential undersampling of these peaks by IM (see below), this is a somewhat artificial consequence of the manner in which the data are compiled (because individual TWIMS separations are certainly fast enough).

Furthermore, factors that contribute to IM peak broadening were investigated. It is well-documented that diffusional band broadening is predominantly responsible for the increased peak widths observed for larger ions in IM [45]. This is evident from a plot of measured IM peak widths for selected poly-DL-alanine ions as a function of scan range and IM cycle time presented in **Figure 6.5** (compare the same coloured spots for a different $t_{d,cycle}$ times, red arrow). More crucial is the peak broadening indicated by the black arrow in this figure, which is a consequence of the manner in which the instrument acquires the data. Because each data point in the mobilogram corresponds to a TOF push event, and the pusher frequency decreases with the upper mass of the scan range (**Table S-3**), the time between individual data points increases for higher scan ranges. As the arrival time of a particular ion remains constant, fewer data points are recorded across each IM peak for higher mass ranges. This is evident from the inset in **Figure 6.5**, where the measured peak width for ion m/z 372 is depicted for four different cycle times (scan ranges). Thus, the inherent feature of the instrument to acquire the same number of data points across each mobilogram leads to some additional broadening of especially low mass ions, and a corresponding loss in resolution or peak capacity when analyzing molecules of high molar mass. The effect of this on the experimentally determined CCS values was, however, relatively small (RSD of 0.03−1.06% (average 0.5%) for n = 9 ions, m/z 300−1300, measured across four scan ranges).

6.3.3 Evaluation of the separation performance of the 3D HILIC×RP-LC×IM system

In order to quantitatively estimate the performance of the HILIC×RP-LC×IM−MS separations, 3D peak capacities were calculated according to established theory, taking both undersampling and orthogonality into account. **Table 6.1** provides a summary of the values obtained; for more detailed information the reader is referred to **Section S-2** and **Table S-4**. (Note that in the evaluation of separation performance, MS is not considered as an additional separation step but rather as the detection step; in principle the following values may be adapted to reflect the additional separation according to m/z as reported previously [21]. We did not do so in light of common practice, where LC−MS is not generally considered a 2D, i.e., LC×MS, system.)

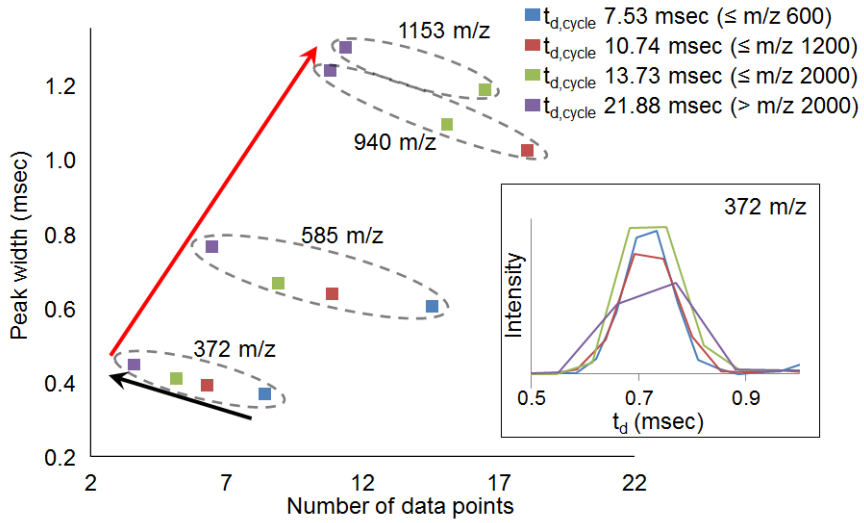


Figure 6.5: Illustration of the relationship between the IM peak width at baseline, m/z (red arrow), and upper mass range (black arrow, also related to the ion mobility cycle time, $t_{d,cycle}$) for selected poly-DL-alanine ions. Colours represent different IM cycle times corresponding to various upper mass limits as defined in the top right corner. The inset shows the experimental peak width for the m/z 372 ion across four different cycle times (scan ranges).

Because a ^{1}D detector was not used in the present work to avoid additional band-broadening, the practical ^{1}D HILIC peak capacity was determined on the basis of the

number of slices detected for each peak and the relevant modulation periods; this value therefore accounts for undersampling of ^{1}D peaks. The ^{2}D peak capacity was determined using the average measured peak widths and the relevant gradient times [46]. Using these values, an average undersampling-corrected peak capacity of 2344 was obtained for the HILIC×RP-LC separations (**Table S-4**). IM peak capacity was determined according to Grushka [47,48], providing an average value of 35 for the four samples analysed here. To account for undersampling of the ^{2}D peaks as a result of IM−MS detection, the scan time of the MS (0.2 s with a 0.024 s interscan delay) was used to calculate the undersampling-correction factor, $^{2}\beta$ [49] (the cycle time of IM separation, 13.7 ms, was not used because of the summation of mobilograms, as discussed above). An average value of 1.30 for $^{2}\beta$ determined in this manner confirms evident undersampling for ^{2}D peaks due to the relatively long scan time, resulting in an average 23% reduction in the practical 3D peak capacity, $n_{c,3D}$, corrected for undersampling (**Table S-4**). Nevertheless, an average $n_{c,3D}$ value of 80 937 was obtained for the HILIC×RP-LC×IM system, confirming the potential of the technique for high-resolution separations.

This value does not present a complete picture though, because correlation between the separation dimensions should be taken into account to reflect the fact that the entire 3D separation space is not accessible to all compounds. In order to account for the finite orthogonality of the three separations, the asterisk method [50] was used. This approach uses an asterisk number (A_0) to express the spreading of the peaks around the separation space as a percentage (A_0 = 100% represents full orthogonality). The average A_0 values for the HILIC×RP-LC ($[A_0]_{12}$), HILIC×IM ($[A_0]_{13}$), and RP-LC×IM ($[A_0]_{23}$) combinations were determined to be 67, 52, and 73%, respectively, confirming good orthogonality (**Table S-4**). The lower orthogonality of HILIC and IM likely reflects the fact that HILIC retention increases with MW for the studied phenolics [33]. (Evaluation of the orthogonality of IM and MS separations provided an expected low average A_0 value of 28%, very similar to the 29% reported for trapped IM−MS analysis of lipids [21]). Despite the good orthogonality measured for any two of the separation dimensions, combination of all three provided a much lower overall value of $[A_0]_{123}$ = 25%. This can be explained by the fact that incomplete orthogonality of any 2D separation results in distribution of analyte peaks across only a portion of the 2D space, which limits the accessible portion of the 3D space upon

addition of a third separation (further subject to the orthogonality considerations of the ^{3}D separation).

Because the asterisk number does not provide a direct measure of the surface coverage, it cannot be used to calculate effective peak capacity, but rather it provides an estimate of the separation power by determining the product of peak capacities (corrected for undersampling) in each dimension and $[A_0]_{123}$ [21]. Values for the separation power calculated in this manner (see **Table 6.1**) vary between 15 974 and 30 072. While much lower than the undersampling-corrected practical peak capacities, this still represents a significant performance gain (on average 13 times higher separation performance) compared to HILIC×RP-LC (**Table S-4**). In **Section S-2** in the SI, orthogonality data determined using the convex-hull method [51] modified for three dimensions are presented, together with practical peak capacities determined using this measure of fractional surface coverage. From the data summarised in **Table S-4**, it can be concluded that in general, similar orthogonality and performance values are obtained using these complementary methods, and therefore, the conclusions discussed above for the asterisk method indeed reflect the experimental reality. It should also be noted that because these values were obtained for diverse classes of phenolic compounds (hydrolysable and condensed tannins, flavonoids, and phenolic acids), the assessment presented above can be considered generally applicable to the analysis of phenolics by HILIC×RP-LC×IM−MS.

Table 6.1: Performance parameters for the comprehensive 3D HILIC×RP-LC×IM−MS separation of phenolics in four different samples.

Performance parameters	Chestnut	Wine	Tea	Grape seed	Avg.
[a] HILIC practical peak capacity, ${}^{1}n'_{c}$	58	46	56	72	58
[b] RP-LC peak capacity, ${}^{2}n_{c}$	29	47	49	38	41
[c] ^{2}D undersampling factor, ${}^{2}\beta$	1.25	1.35	1.39	1.21	1.30
[d] IM peak capacity, ${}^{3}n_{c}$	36	36	34	33	35
[e] Degree of orthogonality, $[A_0]_{123}$ (%)	26	24	17	33	25
[f] Practical peak capacity, $n'_{c,3D}$	61438	77203	93979	91128	80937
[g] Separation power	15974	18529	15976	30072	20138

[a] Calculated as discussed in the text
[b] Calculated according to [46]
[c] Calculated according to [49] using a ^{3}D cycle time of 0.224 sec.
[d] Calculated according to [47]
[e] Calculated according to [50]
[f] $n'_{c,3D} = {}^{1}n'_{c} \times {}^{2}n'_{c} \times {}^{3}n_{c}$
[g] Indicative separation power, corrected for undersampling [49] and orthogonality [50]

6.4 Conclusions

An evaluation of the utility of ion mobility spectrometry in a comprehensive three-dimensional HILIC×RP-LC×TWIM−MS setup revealed several benefits and limitations of the current technology. Important advantages include the availability of drift-time-filtered mass-spectral data for improved compound characterisation, separation of isomers not distinguishable by LC×LC or HR-MS, and increased MS sensitivity. The lack of suitable commercial software for representation and interpretation of the obtained 4D data was addressed by developing an approach for data extraction and presentation in an interactive 3D graph format. For the methodology to gain more widespread application, however, it is imperative that implementation of the data format in automated peak-annotation software protocols be realised. Instrument-related limitations include the software summation of mobilograms corresponding to the MS scan time, which results in some undersampling of narrow ^{2}D RP-LC peaks, as well as under-sampling of mobilogram peaks as a consequence of the relationship between IM cycle time and the mass scan range. The detrimental effect of higher mass scan ranges on the performance of the 3D system points to potential limitations for

high molar mass compounds such as biomolecules, although it should be noted that these observations are confined to the instrumentation used in the present work. It would therefore be informative to evaluate different IM instruments in this context. Nevertheless, our data demonstrate an average increase in the practical separation power of HILIC×RP-LC−MS systems by a factor of 13 by the addition of IM, which clearly highlights the benefits of the approach for complex sample analysis. Considering both the relative separation mechanisms and performance of LC×LC and IM, coupled with the resolving power of HR-MS, it should be noted that IM will only in rare instances provide separation of species not distinguishable by LC×LC−MS. Rather, the benefit of incorporating IM into such systems is expected to primarily be in terms of improving mass-spectral data, providing an additional identification criterion in terms of arrival time, and providing potentially useful information on gas phase conformations.

6.5 References

1. Davis JM, Giddings JC (1983) Statistical-Theory of Component Overlap in Multicomponent Chromatograms. Anal Chem 55:418–424. doi: 10.1021/ac00254a003

2. Regnier F, Amini A, Chakraborty A, Geng M, Ji J, Riggs L, Sioma C, Wang S, Zhang X (2001) Multidimensional Chromatography and the Signature Peptide Approach to Proteomics. LCGC 19:200–213

3. Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. Mol Cell proteomics 1:845–867. doi: 10.1074/mcp.R200007-MCP200

4. Schoenmakers PJ, Vivo G, Decrop WMC (2006) A protocol for designing comprehensive two-dimensional liquid chromatography separation systems. J Chromatogr A 1120:282–290. doi: 10.1016/j.chroma.2005.11.039

5. Guiochon G, Marchetti N, Mriziq K, Shalliker RA (2008) Implementations of two-dimensional liquid chromatography. J Chromatogr A 1189:109–168. doi: 10.1016/j.chroma.2008.01.086

6. Davydova E, Schoenmakers PJ, Vivó-truyols G (2013) Study on the performance of different types of three-dimensional chromatographic systems. J Chromatogr A 1271:137–143. doi: 10.1016/j.chroma.2012.11.043

7. Wouters B, Davydova E, Wouters S, Vivo-Truyols G, Schoenmakers PJ, Eeltink S (2015) Towards ultra-high peak capacities and peak-production rates using spatial three-dimensional liquid chromatography. Lab Chip 15:4415–4422. doi: 10.1039/C5LC01169H

8. Moore AW, Jorgenson JW (1995) Comprehensive Three-Dimensional Separation of Peptides Using Size Exclusion Chromatography/Reversed Phase Liquid Chromatography/Optically Gated Capillary Zone Electrophoresis. Anal Chem 67:3456–3463. doi: 10.1021/ac00115a014

9. Kanu AB, Dwivedi P, Tam M, Matz L, Herbert H (2008) Ion mobility – mass spectrometry. J Mass Spectrom 43:1–22. doi: 10.1002/jms

10. D'Atri V, Causon T, Hernandez-Alba O, Mutabazi A, Veuthey JL, Cianferani S, Guillarme D (2018) Adding a new separation dimension to MS and LC–MS: What is the utility of ion mobility spectrometry? J Sep Sci 41:20–67. doi: 10.1002/jssc.201700919

11. Pringle SD, Giles K, Wildgoose JL, Williams JP, Slade SE, Thalassinos K, Bateman RH, Bowers MT, Scrivens JH (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. Int J Mass Spectrom 261:1–12. doi: 10.1016/j.ijms.2006.07.021

12. Holcapek M, Jirásko R, Lísa M (2012) Recent developments in liquid chromatography – mass spectrometry and related techniques. J Chromatogr A 1259:3–15. doi: 10.1016/j.chroma.2012.08.072

13. Zheng X, Renslow RS, Makola MM, Webb IK, Deng L, Thomas DG, Govind N, Ibrahim YM, Kabanda MM, Dubery IA, Heyman HM, Smith RD, Madala NE, Baker ES (2017) Structural Elucidation of cis/trans Dicaffeoylquinic Acid Photoisomerization Using Ion Mobility Spectrometry-Mass Spectrometry. J Phys Chemsitry Lett 8:1381–1388

14. Lanucara F, Holman SW, Gray CJ, Eyers CE (2014) The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. Nat Chem 6:281–294. doi: 10.1038/nchem.1889

15. Paglia G, Williams JP, Menikarachchi L, Thompson JW, Tyldesley-Worster R, Halldórsson S, Rolfsson O, Moseley A, Grant D, Langridge J, Palsson BO, Astarita G (2014) Ion mobility derived collision cross sections to support metabolomics applications. Anal Chem 86:3985–3993. doi: 10.1021/ac500405x

16. Dodds JN, May JC, McLean JA (2017) Correlating Resolving Power, Resolution, and Collision Cross Section: Unifying Cross-Platform Assessment of Separation Efficiency in Ion Mobility Spectrometry. Anal Chem 89:12176–12184. doi: 10.1021/acs.analchem.7b02827

17. Stephan S, Jakob C, Hippler J, Schmitz OJ (2016) A novel four-dimensional analytical approach for analysis of complex samples. Anal Bioanal Chem

408:3751–3759. doi: 10.1007/s00216-016-9460-9

18. Stephan S, Hippler J, Köhler T, Deeb AA, Schmidt TC, Schmitz OJ (2016) Contaminant screening of wastewater with HPLC-IM-qTOF-MS and LC+LC-IM-qTOF-MS using a CCS database. Anal Bioanal Chem 408:6545–6555. doi: 10.1007/s00216-016-9820-5

19. Ehkirch A, D'Atri V, Rouviere F, Hernandez-Alba O, Goyon A, Colas O, Sarrut M, Beck A, Guillarme D, Heinisch S, Cianferani S (2018) An Online Four-Dimensional HIC×SEC-IM×MS Methodology for Proof-of-Concept Characterization of Antibody Drug Conjugates. Anal Chem 90:1578–1586. doi: 10.1021/acs.analchem.7b02110

20. Marriott PJ, Schoenmakers P, Wu Z (2012) Nomenclature and Conventions in Comprehensive Multidimensional Chromatography - An Update. LCGC Eur 25:266–275

21. Baglai A, Gargano AFG, Jordens J, Mengerink Y, Honing M, van der Wal S, Schoenmakers PJ (2017) Comprehensive lipidomic analysis of human plasma using multidimensional liquid- and gas-phase separations: Two-dimensional liquid chromatography–mass spectrometry vs. liquid chromatography–trapped-ion-mobility–mass spectrometry. J Chromatogr A 1530:90–103. doi: 10.1016/j.chroma.2017.11.014

22. Stander MA, Van Wyk B, Taylor MJC, Long HS (2017) Analysis of Phenolic Compounds in Rooibos Tea (Aspalathus linearis) with a Comparison of Flavonoid-Based Compounds in Natural Populations of Plants from Different Regions. J Agric Food Chem 65:10270–10281. doi: 10.1021/acs.jafc.7b03942

23. Moss R, Mao Q, Taylor D, Saucier C (2013) Investigation of monomeric and oligomeric wine stilbenoids in red wines by ultra-high-performance liquid chromatography / electrospray ionization quadrupole time-of- flight mass spectrometry. 1815–1827. doi: 10.1002/rcm.6636

24. Kalili KM, de Villiers A (2009) Off-line comprehensive 2-dimensional hydrophilic interaction × reversed phase liquid chromatography analysis of procyanidins. J Chromatogr A 1216:6274–6284. doi: 10.1016/j.chroma.2009.06.071

25. Ruotolo BT, Benesch JL, Sandercock AM, Hyung S, Robinson C V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes. Nat Protoc 3:1139–1152

26. Chambers MC, MacLean B, Burke R, Amode D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak M-Y, Paulse C, Creasy D, Flashner L, Kani K, Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J, Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M, Tabb DL, Mallick P (2012) A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol 30:918–920. doi: 10.1038/nbt.2377

27. Python Software Foundation. Python Language Reference, Version 3.6. Python Software Foundation.

28. Pérez F (2007) Granger, B.E. IPython: A System for Interactive Scientific Computing. Comput Sci Eng 9:21–29

29. Goloborodko AA, Levitsky LI, Ivanov M V, Gorshkov M V (2013) Pyteomics — a Python Framework for Exploratory Data Analysis and Rapid Software Prototyping in Proteomics. J Am Soc Mass Spectrom 24:301–304. doi: 10.1007/s13361-012-0516-6

30. McKinney W (2010) Data Structures for Statistical Computing in Python. In: Proceedings of the 9th Python in Science Conference. pp 51–56

31. Plotly Technologies Inc. (2015) Collaborative data science

32. Cacciola F, Farnetti S, Dugo P, Marriott PJ, Mondello L (2016) Comprehensive Two-dimensional Liquid Chromatography for Polyphenol Analysis in Foodstuffs. J Sep Sci In press:. doi: 10.1002/jssc.201600704.This

33. De Villiers A, Kalili KM (2017) Comprehensive Two-Dimensional Hydrophilic Interaction Chromatography × Reversed Phase Liquid Chromatography (HILIC×RP–LC): Theory, Practice, and Applications. In: Grushka E, Grinberg N (eds) Advances in Chromatography, Vol. 53. CRC Press,Taylor & Francis Group, pp 217–299

34. Muller M, Tredoux AGJ, de Villiers A (2018) Predictive kinetic optimisation of hydrophilic interaction chromatography × reversed phase liquid chromatography separations: Experimental verification and application to phenolic analysis. J Chromatogr A 1571:107–120. doi: 10.1016/j.chroma.2018.08.004

35. Stoll DR, Talus ES, Harmes DC (2015) Evaluation of detection sensitivity in comprehensive two-dimensional liquid chromatography separations of an active pharmaceutical ingredient and its degradants. Anal Bioanal Chem 407:265–277. doi: 10.1007/s00216-014-8036-9

36. Muller, M. Strategies for method development in on-line comprehensive 2-dimensional liquid chromatography. MSc thesis 2018, Stellebosch University http://scholar.sun.ac.za/.

37. Gonzales GB, Smagghe G, Coelus S, Adriaenssens D, De Winter K, Desmet T, Raes K, Van Camp J (2016) Collision cross section prediction of deprotonated phenolics in a travelling-wave ion mobility spectrometer using molecular descriptors and chemometrics. Anal Chim Acta 924:68–76. doi: 10.1016/j.aca.2016.04.020

38. Sanz M, Cadahia E, Esteruelas E, Munoz AM, Fernandez De Simon B, Hernandez T, Estrella I (2010) Phenolic compounds in chestnut (Castanea sativa Mill.) heartwood. Effect of toasting at cooperage. J Agric Food Chem 58:9631–9640. doi: 10.1021/jf102718t

39. Monagas M, Bartolomé B, Gómez-Cordovés C (2005) Updated knowledge about the presence of phenolic compounds in wine. Crit Rev Food Sci Nutr 45:85–118. doi: 10.1080/10408690490911710

40. Kalili KM, Vestner J, Stander MA (2013) Toward Unraveling Grape Tannin Composition: Application of Online Hydrophilic Interaction Chromatography × Reversed-Phase Liquid Chromatography − Time-of-Flight Mass Spectrometry for Grape Seed Analysis. Anal Chem 85:9107–9115

41. O'Kennedy, S. J.; Gerber, W. J.; de Villiers, A.; Brand, D. J. A Variable Temperature 1 H NMR and DFT Study of Procyanidin B2 Conformational Interchange. Struct. Chem. 2018, doi.org/10.1007/s11224-018-1153-x

42. Kuhnert N, Yassin GH, Jaiswal R, Matei MF, Grün CH (2015) Differentiation of prototropic ions in regioisomeric caffeoyl quinic acids by electrospray ion mobility mass spectrometry. Rapid Commun Mass Spectrom 29:675–680

43. Yassin GH, Grun C, Koek JH, Assaf I, Kuhnert N (2014) Investigation of isomeric flavanol structures in black tea thearubigins using ultraperformance liquid chromatography coupled to hybrid quadrupole/ion mobility/time of flight mass spectrometry. J Mass Spectrom 1086–1095. doi: 10.1002/jms.3406

44. Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH (2004) Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. Rapid Commun Mass Spectrom 18:2401–2414. doi: 10.1002/rcm.1641

45. Shvartsburg AA, Smith RD (2008) Fundamentals of Traveling Wave Ion Mobility Spectrometry. Anal Chem 80:9689–9699

46. Neue UD (2005) Theory of peak capacity in gradient elution. J Chromatogr A 1079:153–161. doi: 10.1016/j.chroma.2005.03.008

47. Grushka E (1970) Chromatographic Peak Capacity and the Factors Influencing It. Anal Chem 42:1142–1147. doi: 10.1021/ac60293a001

48. Causon TJ, Hann S (2015) Theoretical evaluation of peak capacity improvements by use of liquid chromatography combined with drift tube ion mobility-mass spectrometry. J Chromatogr A 1416:47–56. doi: 10.1016/j.chroma.2015.09.009

49. Li X, Stoll DR, Carr PW (2009) Equation for peak capacity estimation in two-dimensional liquid chromatography. Anal Chem 81:845–850. doi: 10.1021/ac801772u

50. Camenzuli M, Schoenmakers PJ (2014) A new measure of orthogonality for multi-dimensional chromatography. Anal Chim Acta 838:93–101. doi: 10.1016/j.aca.2014.05.048

51. Semard G, Peulon-Agasse V, Bruchet A, Bouillon JP, Cardinaël P (2010) Convex hull: A new method to determine the separation space used and to optimize operating conditions for comprehensive two-dimensional gas

chromatography. J Chromatogr A 1217:5449–5454. doi: 10.1016/j.chroma.2010.06.048

Chapter 6

Supporting information for: Comprehensive 3-dimensional LC×LC×ion mobility spectrometry separation combined with high resolution MS for the analysis of complex samples

Table S-1: Chromatographic conditions used for the on-line HILIC×RP-LC separation of phenolics in four natural product samples.

Sample	**Chestnut**		**Grape seed**	
Dimension	^{1}D	^{2}D	^{1}D	^{2}D
Flow rate	9 µL/min	3.0 mL/min	11 µL/min	3.0 mL/min
Dilution flow (µL/min)	126		99	
Modulation time (min)	0.47		0.58	
Gradient time (min)	0.27		0.38	
Injection volume (µL)	1	63	2	64
Injection solvent	50/50 MeOH/H_2O		50/50 MeOH/ACN	
Gradient (MP A: 0.1% FA and MP B: ACN)	10-20% A in 0-20 min;20-50% A in 20-30 min; 50-98% A in 35-50 min; 98% A 50-60 min	1-15%B in 0-0.15 min; 15-45% B in 0.15-0.27 min; 45-100% in 0.27-0.28 min	2-10% A in 0-5 min; 10-35% in 5-45 min; 35-90% in 45-60 min; 90% to 70 min	1-40% B in 0-0.38min; 40-100% B 0.38-0.39 min

Sample	Wine		Tea	
Dimension	^{1}D	^{2}D	^{1}D	^{2}D
Flow rate	11 µL/min	2.6 mL/min	11 µL/min	2.6 mL/min
Dilution flow (µL/min)	99		99	
Modulation time (min)	0.58		0.58	
Gradient time	0.35		0.35	
Injection volume (µL)	2	64	2	64
Injection solvent	25/75 MeOH/ACN		50/50 MeOH/ACN	
Gradient (MP A: 0.1% FA and MP B: ACN)	2-27% A in 0-45 min; 27-90% in 45-60 min; 90 % to 70 min	1-45% B in 0-0.35min; 45-100% B 0.35-0.36 min	2-10% A in 0-5 min; 10-35% in 5-45 min; 35-90% in 45-60 min; 90% to 70 min	1-45% B in 0-0.35 min; 45-100% B in 0.35-0.36 min

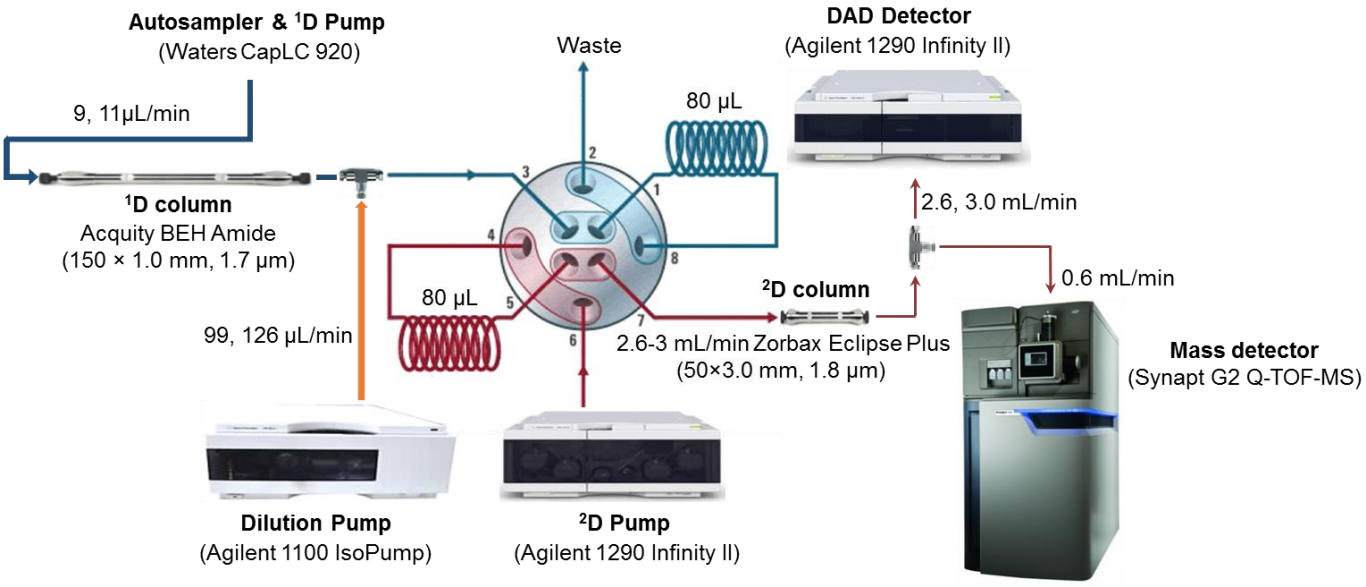


Figure S-1: Schematic illustration of the instrumental setup and conditions used for the HILIC×RP-LC separation coupled to a DAD and IM-Q-TOF detection. The ^{2}D flow (**Table S-1**) was split 4:1 between the DAD and MS detectors post-column.

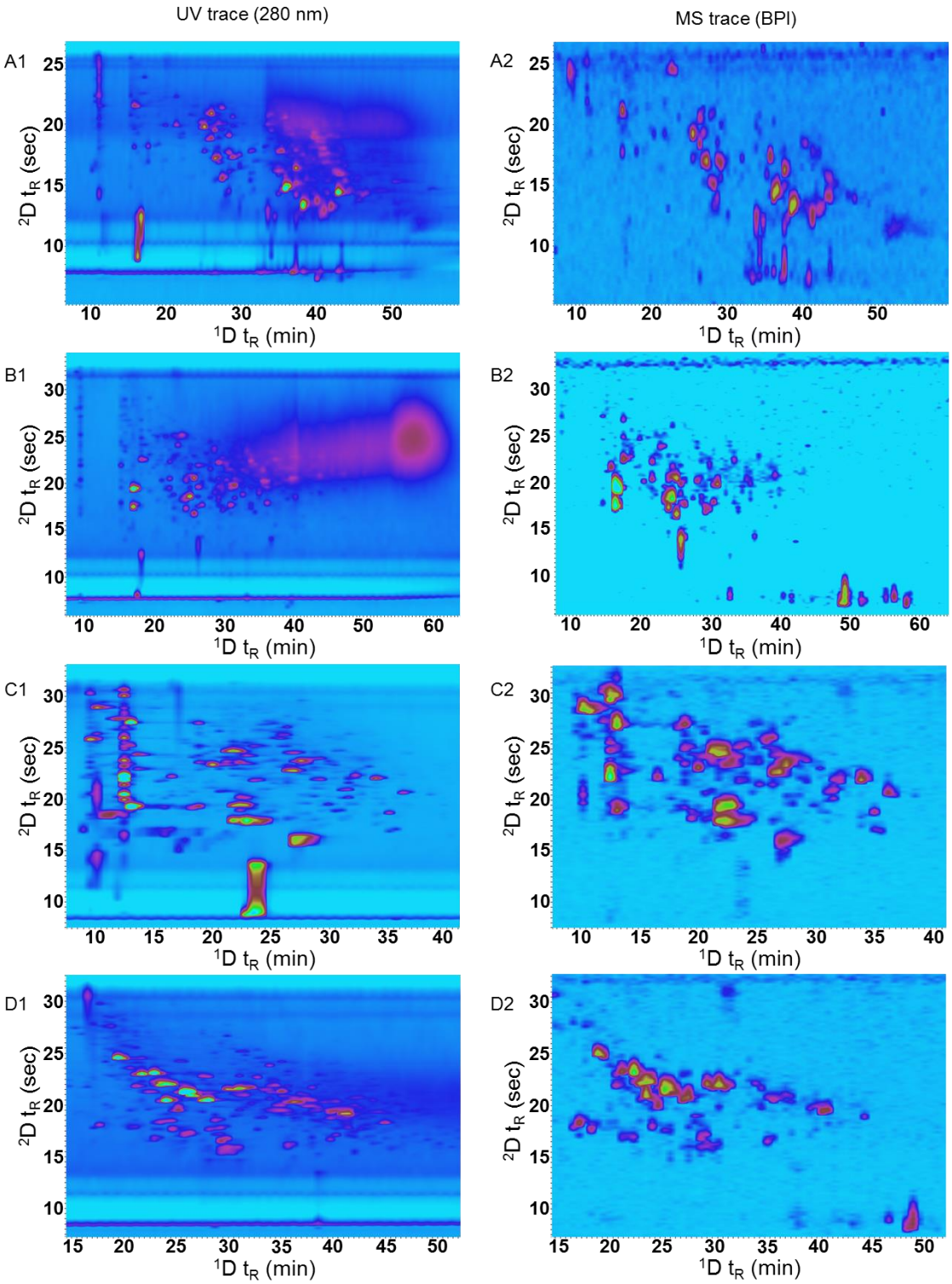


Figure S-2: Two-dimensional UV (280 nm, left) and MS (BPI, right) contour plots obtained for the on-line HILIC×RP-LC separation of (**A**) chestnut, (**B**) grape seed, (**C**) wine and (**D**) rooibos tea phenolics. For experimental conditions, refer to **Table S-1**.

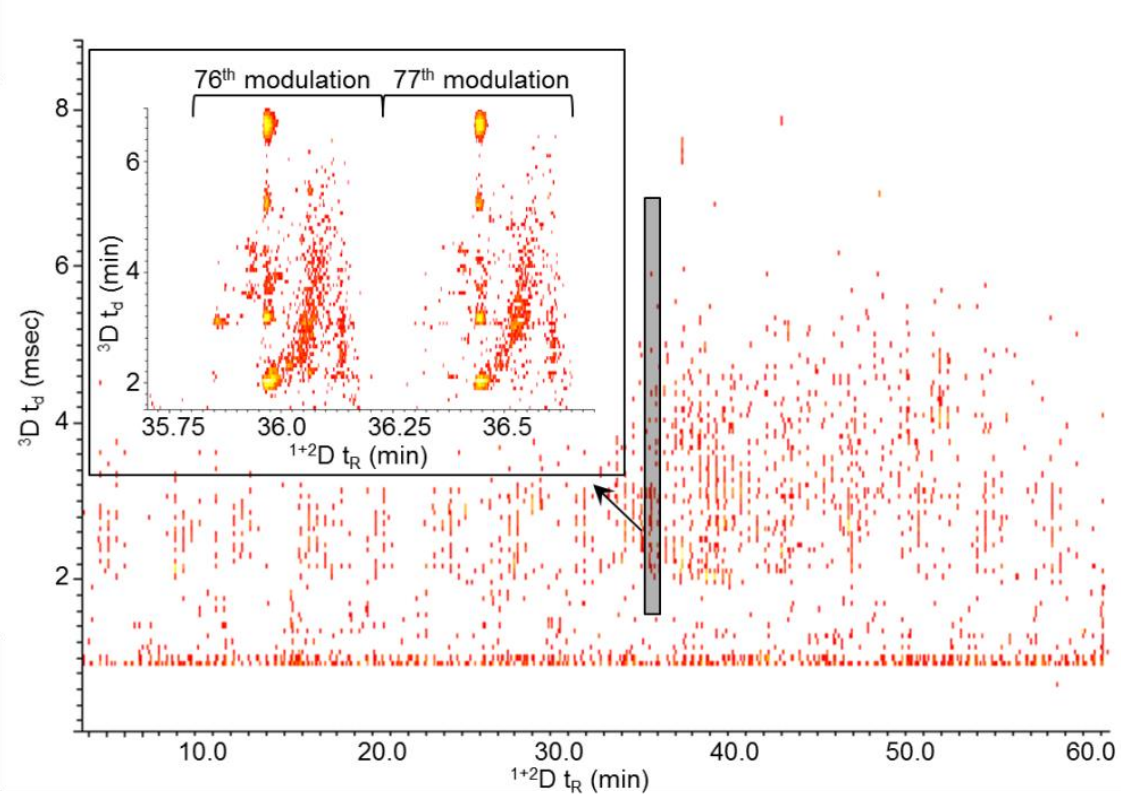


Figure S-3: Example of the raw mobilogram data obtained for the on-line HILIC×RP-LC×IM-MS separation of chestnut tannins. The insert shows two consecutive ^{2}D modulation periods of 0.47 min with the corresponding two RP-LC gradient separations. This representation was obtained using Driftscope software (Waters).

Table S-2: Examples of the identification of selected phenolic compounds in chestnut, rooibos tea, grape seed and wine samples based on retention times in both HILIC and RP-LC separations, UV and HR-MS spectral data and measured CCS values. Experimental CCS values were obtained from calibration data for poly-DL-alanine according to [25] and compared to literature reports, where available, or to values obtained on a DTIMS instrument.

Sample	Compound	UV	^{12}C $[M-H]^{1-}$	ppm	t_R HILIC (min)	t_R RP-LC (sec)	t_d (msec)	CCS	CCS_{ref}*	Fragment	Ref
Chestnut	Vescalagin	280 sh	933.0633	0.01	38.07	13.34	6.55	287	290[a]	915, 871, 613, 569, 300, 275, 249	[38]
Rooibos tea	Aspalathin	288	451.1268	0.62	24.94	18.60	2.83	196	192[b]	361, 331	[22]
Grape seed	B-type procyandin dimer	280	577.1345	0.02	25.52	21.18	3.73	223	222[c]	449, 425, 407, 289, 125	[40]
Wine	Myricetin	370	317.0295	0.06	21.46	24.68	1.93	163	160[d], 168[e]	179, 151, 137	[39]

[sh] indicates the presence of a shoulder.
*CCS_{ref} values derived from literature reports (TWIMS) and determined on a drift tube (DT) IMS spectrometer as specified below.
[a,c and e]CCS values determined on a drift tube (DT) IM-MS (Agilent 6560 IMS-QTOFMS), Department of Chemistry, BOKU Vienna, Austria.
[b]Stander et. al. *J. Agric. Food Chem.* 2017, *65* (47), 10270–10281.
[d]Gonzales et. al. *Anal. Chim. Acta* 2016, *924*, 68–76.

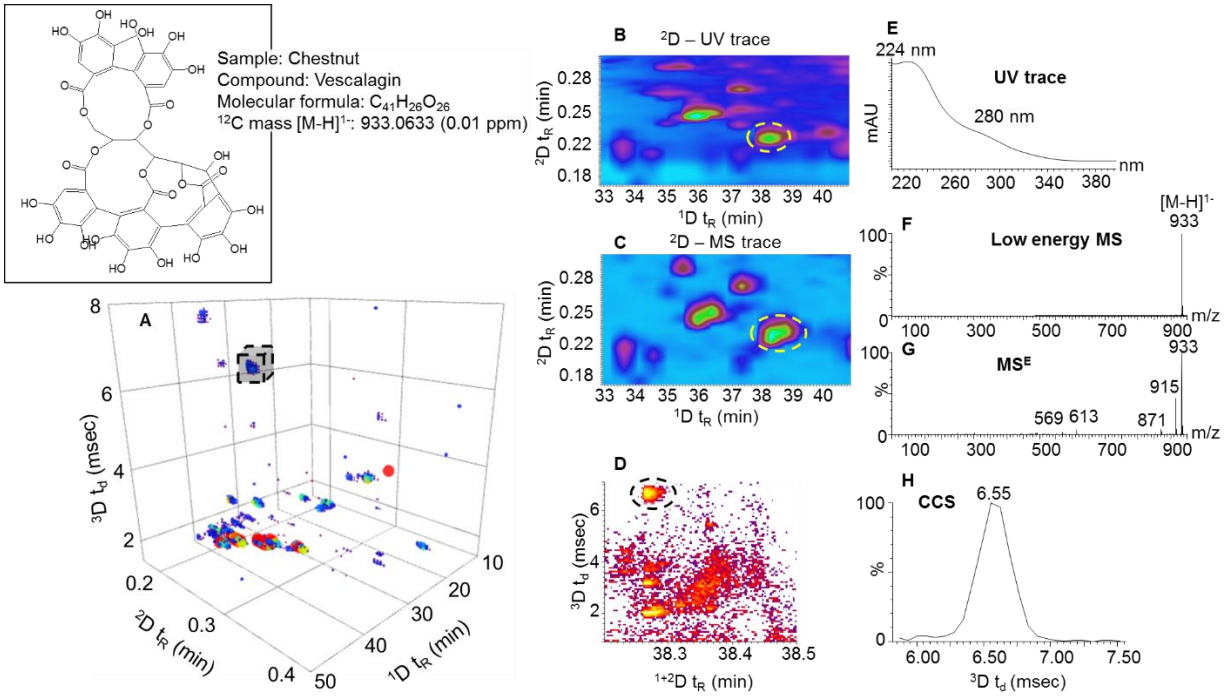


Figure S-4: Illustration of the identification of Vescalagin in the chestnut sample based on HILIC×RP-LC×IM-MS data. (**A**) Shows the three-dimensional representation of the data, (**B**) the UV, (**C**) TIC and (**D**) arrival time contour plots of the relevant separation window. (**E**) UV, (**F**) low- and (**G**) high collision energy mass spectra and (**H**) the extracted ion mobilogram for Vescalagin are presented on the right. The position of the Vescalagin peak is indicated by the dotted black square and yellow/black circles. For further details on this compound, refer to **Table S-2**.

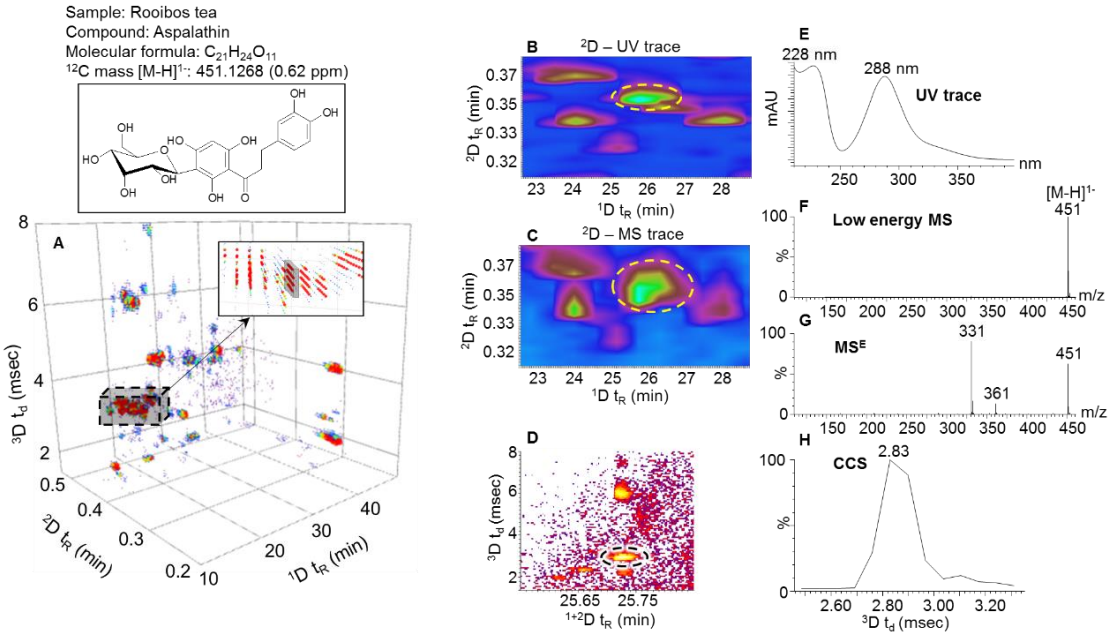


Figure S-5: Illustration of the identification of aspalathin in a rooibos tea sample based on HILIC×RP-LC(×UV)×IM-MS data. (**A**) shows the three-dimensional representation of the data with an insert showing an enlarged area of interest, (**B**) the UV, (**C**) TIC and (**D**) arrival time contour plots of the relevant separation window. (**E**) UV, (**F**) low- and (**G**) high collision energy mass spectra and (**H**) the extracted ion mobilogram for aspalathin are presented on the right. The position of the aspalathin peak is indicated by the dotted black square and yellow/black circles. For further details on this compound, refer to **Table S-2**.

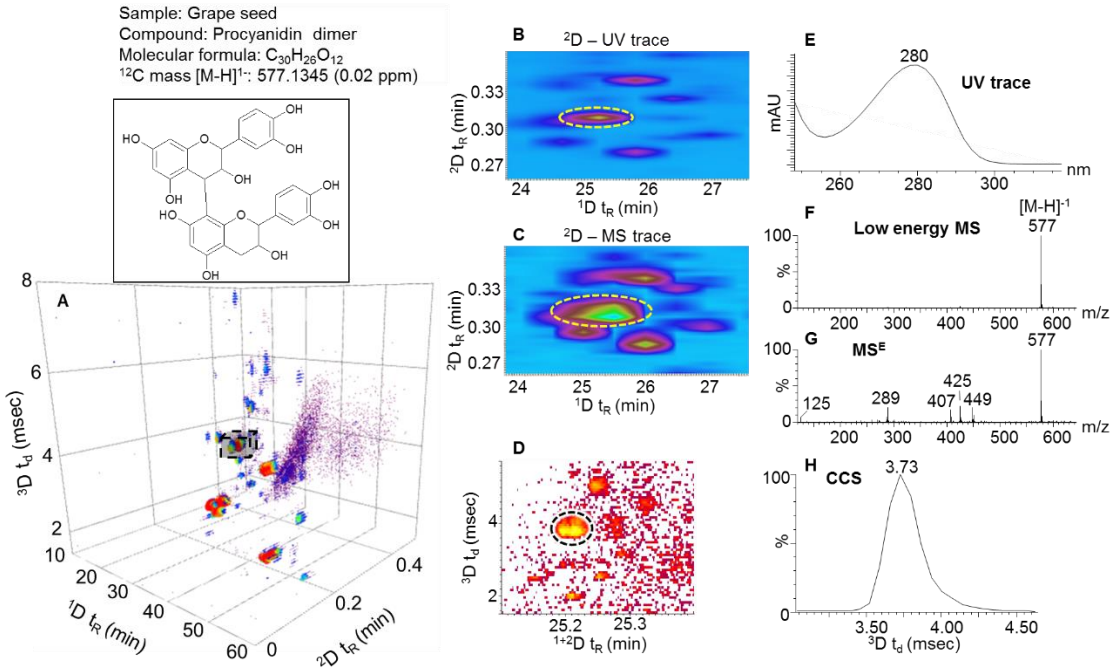


Figure S-6: Illustration of the identification of a procyanidin dimer in a grape seed sample based on HILIC×RP-LC(×UV)×IM-MS data. (**A**) shows the three-dimensional representation of the data, (**B**) the UV, (**C**) TIC and (**D**) arrival time contour plots of the relevant separation window. (**E**) UV, (**F**) low- and (**G**) high collision energy mass spectra and (**H**) the extracted ion mobilogram for a procyanidin dimer are presented on the right. The position of the procyanidin dimer peak is indicated by the dotted black square and yellow/black circles. For further details on this compound, refer to **Table S-2**.

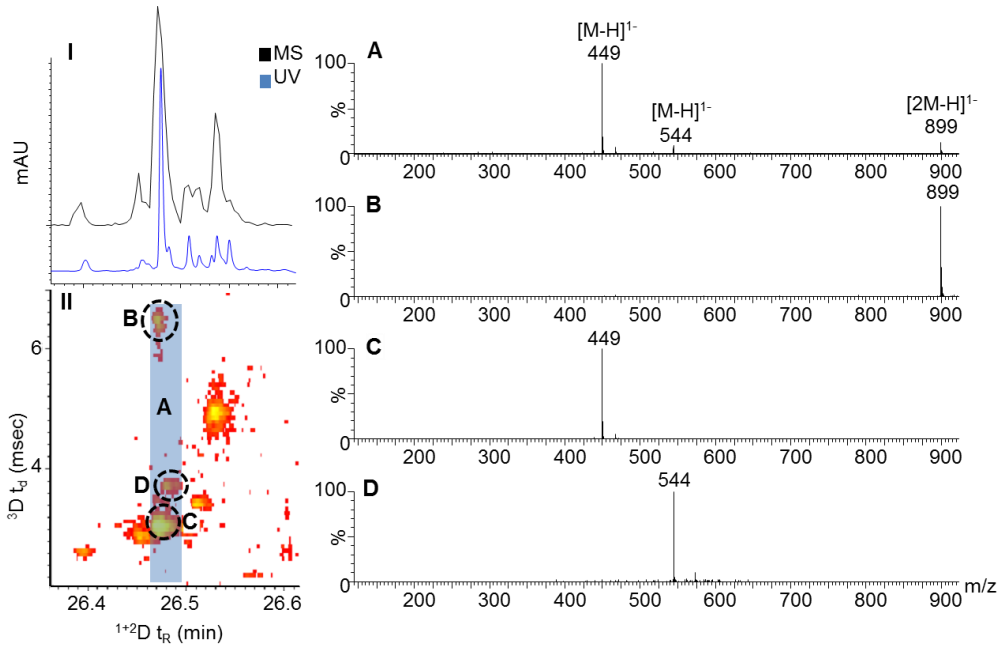


Figure S-7: (Left): A comparison of TIC (black) and UV (blue) traces (**I**) for a single ^{2}D RP-LC separation illustrating the loss in resolution for the MS trace due to extra-column band broadening and the relatively low acquisition rate of the MS detector. (**II**): arrival time contour plot for the same ^{2}D separation, showing multiple arrival times for co-eluting species. (Right): (**A**) mass spectrum of the chromatographic trace (shaded in blue), (**B**), (**C**) and (**D**) show the arrival time-filtered MS spectra for arrival times of 6.42, 2.97 and 3.66 msec, respectively.

Section S-1: Sensitivity as a function of mass scan range

In order to investigate the effect of the scanned mass range on sensitivity, poly-DL-alanine was injected and data acquired for different scan ranges with and without IM activated. **Figure S-8** shows a comparison of mass spectra measured for poly-DL-alanine with and without IM using a mass range of 300-2000 m/z and a scan time of 0.2 sec. The signal intensity# (ion counts) for selected ions* within the respective mass ranges were determined for each scan range, and are summarized for a scan time of 0.2 sec in **Figure S-9**. From these data, significantly elevated ion intensities for IM measurements (red) are evident for mass scan ranges up to m/z 1000 (m/z 372, 585, 727 and 940) compared to MS mode (blue). In addition, signal-to-noise (S/N) ratios were measured for a chromatographic injection of catechin for scanned mass ranges of m/z 50 to 600, 2000 and 5000 both with and without IM activated. A substantial decrease in background noise (≈ 50%) was observed when operating the instrument in ion mobility mode, especially for lower molecular ions (<600 m/z, results not shown). The increased duty cycle observed for IM experiments is due to the trapping and releasing of ions packets during the IM separation and synchronisation of their arrival times with the orthogonal TOF pusher pulses, which results in an enhanced sensitivity for ions within a particular m/z range [44]. For a scan range of 50-1500 m/z, similar signal intensities were measured with and without IM for the ion at m/z 1153. For higher upper mass ranges (≥ m/z 2000), however, much lower signal counts were observed for high mass ions (m/z 1295 and 1487 in **Figure S-9**) with IM activated. This is related to the delay time set between release of ion packets from the IM and TOF pusher. This setting affects (enhances) the duty cycle for a certain range of ions, but not for all. With the instrument settings used in this work, the maximum duty cycle enhancement was observed for ions of m/z 585 and 727, but led to lower sensitivity for higher mass ions. This parameter can be tuned to selectively enhance the intensity of ions within a particular mass range; since most of the compounds detected in the samples analysed in this work were detected at m/z < 1000, this was not attempted here. Although only data for a 0.2 sec scan time are shown, similar trends were observed for all mass ranges for scan times between 0.2 and 0.8 sec (for scan times below 0.2 and above 0.8 seconds, the overall sensitivity decreased).

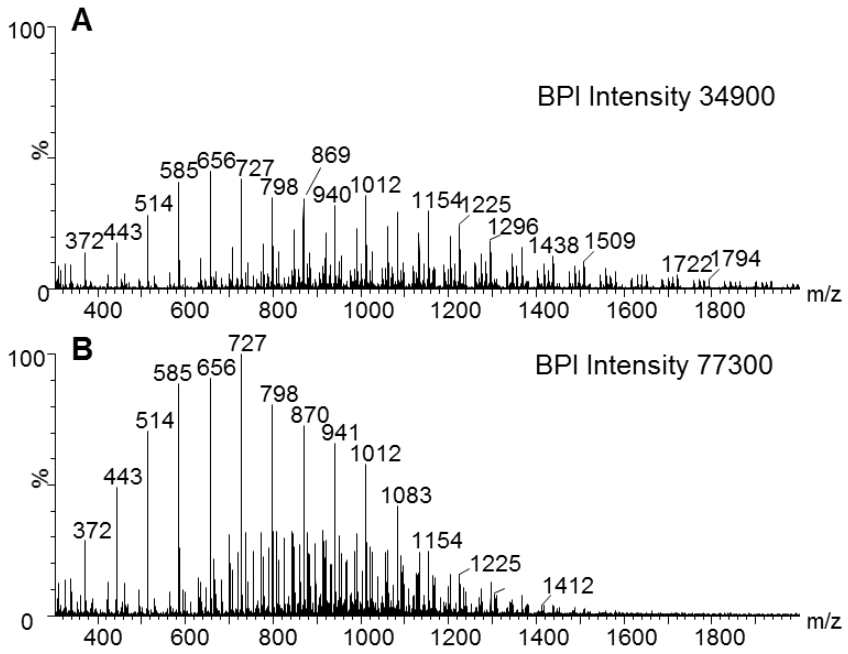


Figure S-8: Examples of infusion mass spectra measured for poly-DL-alanine in negative ionisation mode (**A**) in MS mode, and (**B**) in IM-MS mode. Base peak ion (BPI) intensities are specified for each of these mass spectra. Both spectra acquired for a scan range of 300-2000 m/z and a scan time of 0.2 sec.

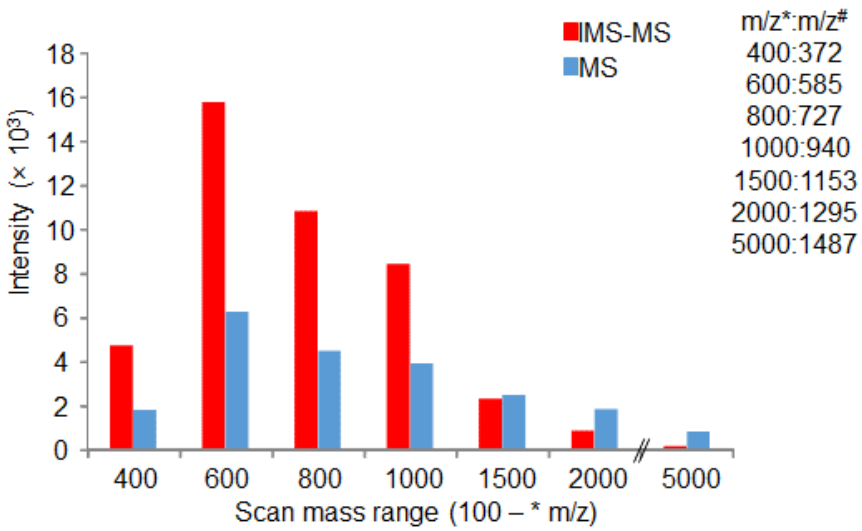


Figure S-9: Comparison of ion intensities as a function of scanned mass range* (100-the specified number, x-axis) for poly-DL-alanine ions with (red) and without (blue) IM. Data are plotted for the ions# within the different mass ranges specified on the right of the graph for a scan time of 0.2 sec.

Table S-3: Relationship between TOF pusher frequency (Hz)[a] and the upper mass limit on the Synapt G2 instrument.

Highest mass scanned (m/z)	TOF pusher frequency Hz[a]	IM cycle time (msec)	Push events/IM cycle time
≤ 600	26455	7.56	200
≤ 1200	18622	10.74	200
≤ 2000	14567	13.73	200
> 2000	9141	21.88	200

[a]Estimated from the IM cycle time and the number of push events per IM cycle time.

Section S-2: Estimation of the practical peak capacity of LC×LC×IM separations

The peak capacities of the two LC separations were calculated using the respective gradient times (t_g) and the average peak widths (4σ) [46]:

$$n = 1 + \frac{t_g}{4\sigma} \quad \text{(S1)}$$

^{1}D peak widths were determined based on the number of slices detected for each peak multiplied by the modulation time. Since the peak width measured in this manner already includes the effect of undersampling, this provides the corrected ^{1}D peak capacity ($^1n'_c$). In the ^{2}D, peak widths at half height were measured for the main peaks in each sample from UV data, and converted to 4σ to calculate the ^{2}D peak capacity (^2n_c). A minimum of 24 compounds spread across the 2D separation space for each sample were used to measure average ^{1}D and ^{2}D peak widths. The practical ^{2}D peak capacity corrected for undersampling ($^2n'_c$) is then calculated using [49]:

$$^2n_c' = \frac{^2n_c}{^2\beta} \quad \text{(S2)}$$

where the ^{2}D undersampling factor, ^{2}β, is defined as

$$^2\beta = \sqrt{1 + 3.35\left(\frac{^3t_c \ ^2n_c}{^2t_g}\right)^2} \quad \text{(S3)}$$

and the ^{3}D cycle time (^3t_c) is equal to the scan time of the MS (in this work 0.224 sec). IM peak widths at half height were measured for the major singly charged peaks in each sample (n = 14-20 compounds per sample, 64 ions of m/z 163-1101 across all samples). These values were used to create a plot of IM peak width at half height vs. arrival time. The peak capacity of the IM separation was then calculated using [47,48]:

$$^{3}n_c = 1 + \frac{\sqrt{5.54}}{4a}\ln\frac{t_{d,n}}{t_{d,1}} \quad \text{(S4)}$$

where *a* is the slope of the IM peak width at half height vs. arrival time plot (r^2 = 0.88-0.91), and $t_{d,1}$ and $t_{d,n}$ represent the arrival time of the first and last peaks, respectively. The practical 2D (HILIC×RP-LC) and 3D (HILIC×RP-LC×IM) peak capacities corrected for undersampling ($n'_{c,2D}$ and $n'_{c,3D}$) were calculated using the Eq.'s S5 and S6, respectively.

$$n'_{c,2D} = {}^{1}n'_c \times {}^{2}n_c \quad \text{(S5)}$$

$$n'_{c,3D} = {}^{1}n'_c \times {}^{2}n'_c \times {}^{3}n_c \quad \text{(S6)}$$

The degree of orthogonality between all three separation dimensions, as well as between IM and MS, were evaluated using both the asterisk [50] and convex hull [51] methods. The asterisk method uses an asterisk number (A_0) to express the spreading of the peaks around the separation space as a percentage (A_0=100% represents full orthogonality). The degree of orthogonality for selected combinations of two dimensions ($[A_0]_{xy}$, where x and y represent the respective dimensions) was calculated in Excel using the equations described by Camenzuli et al.[50]. The degree of orthogonality for the three-dimensional HILIC×RP-LC×IM system was calculated using [21]:

$$[A_0]_{123} = [A_0]_{12} \times [A_0]_{23} \times [A_0]_{13} \quad \text{(S7)}$$

The fractional surface coverage (f_c) between two ($f_{c,xy}$) and three ($f_{c,xyz}$) dimensions (subscripts x, y and z represent the respective dimensions) were calculated using a minimum convex hull area script in Matlab. The range scaled retention times, arrival times and m/z values were used throughout. As an example, the surface coverage plots and orthogonality values obtained using both approaches are presented for the chestnut sample in **Figure S-10**.

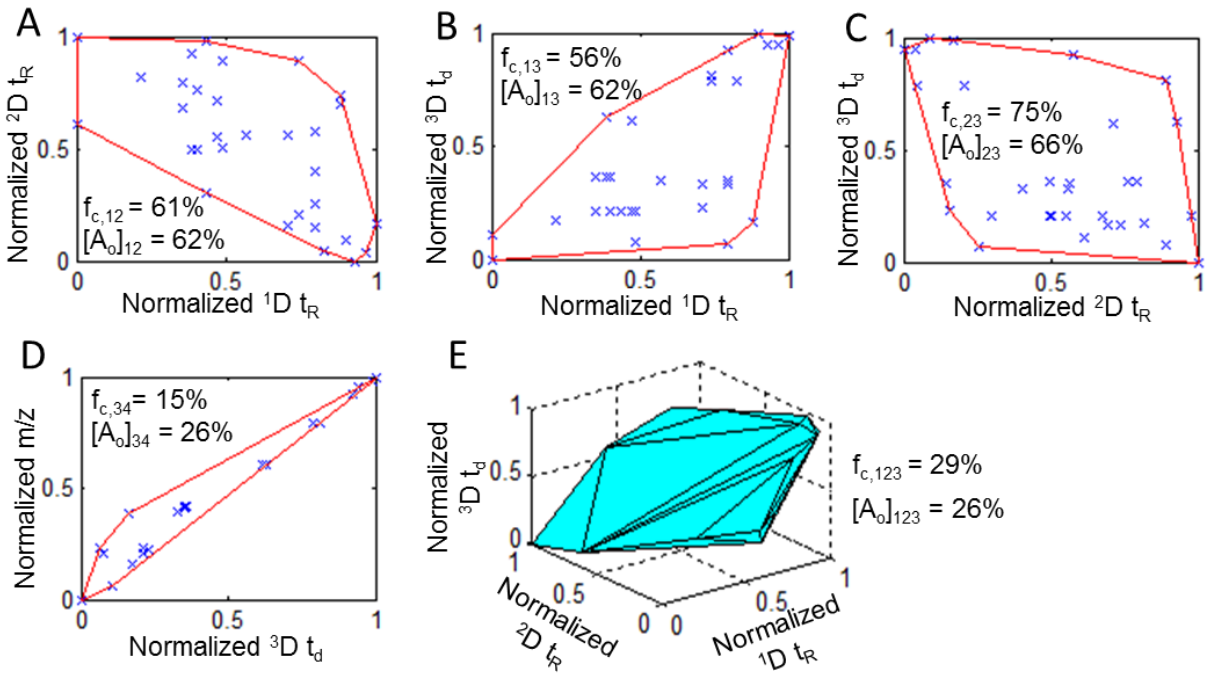


Figure S-10: Graphs illustrating the correlation between the (**A**) HILIC and RP-LC, (**B**) HILIC and IM, (**C**) RP-LC and IM and (**D**) IM and MS separation of chestnut tannins as well as the three dimensional minimum convex hull area of the (**E**) HILIC×RP-LC×IM separation. The degree of orthogonality (A_0) and fractional surface coverage (f_c) values for each combination are indicated on each graph. For the same data on the other samples, the reader is referred to **Table S-4**.

Since the asterisk number does not provide a direct measure of the surface coverage, it cannot be used to calculate effective peak capacities. However, the product of the practical 2D or 3D peak capacity, corrected for undersampling, and the relevant $[A_0]_{12}$ or $[A_0]_{123}$ values provides an estimate of the separation power of the relevant separation system. These values are listed in **Table S-4**.

Furthermore, the undersampling corrected practical 2D and 3D peak capacities ($n'_{c,2D}$ and $n'_{c,3D}$) were further corrected for orthogonality using the fractional surface coverage values ($f_{c,12}$ and $f_{c,123}$) obtained using the convex hull method, to provide the practical peak capacities corrected for both undersampling and orthogonality (denoted $n''_{c,2D}$ and $n''_{c,3D}$, respectively).

$$n''_{c,2D} = n'_{c,2D} \times f_{c,12} \quad \text{(S8)}$$

$$n''_{c,3D} = n'_{c,3D} \times f_{c,123} \quad \text{(S9)}$$

Table S-4: Performance parameters of the 3D (HILIC×RP-LC×IM-MS) separation of the complex polyphenolic samples.

Performance parameters	Chestnut	Wine	Tea	Grape seed	Average
HILIC gradient time (min), ^1t_g	50	60	60	60	58
[a] HILIC average peak width (min)	0.88	1.35	1.09	0.85	1.04
[b] HILIC practical peak capacity, $^1n'_c$	58	46	56	72	58
^{2}D cycle time (min), ^2t_c	0.47	0.58	0.58	0.58	0.55
RP-LC gradient time (sec), ^2t_g	16.2	21	21	22.8	20.3
[c] RP-LC average peak width (sec)	0.57	0.46	0.43	0.61	0.52
[b] RP-LC peak capacity, ^2n_c	29	47	49	38	41
[d] HILIC×RP-LC practical peak capacity, $n'_{c,2D}$	1707	2145	2764	2761	2344
[e] HILIC×RP-LC practical peak capacity, $n''_{c,2D}$	1041	1243	1384	2007	1419
[f] HILIC×RP-LC separation power	1058	1437	1603	2182	1570
^{3}D cycle time (sec), ^3t_c	0.224	0.224	0.224	0.224	0.224
[g] ^{2}D undersampling factor, $^2\beta$	1.25	1.35	1.39	1.21	1.30
[h] RP-LC practical peak capacity, $^2n'_c$	24	34	36	32	32
[c] IM average peak width (msec)	0.40	0.27	0.32	0.40	0.35
[i] IM peak capacity, ^3n_c	36	36	34	33	35
[j] $[A_0]_{12}$: HILIC×RP-LC (%)	62	67	58	79	67
[k] Fractional surface coverage, $f_{c,12}$ (%)	61	58	50	73	60
[j] $[A_0]_{13}$: HILIC×IM (%)	62	55	42	47	52
[k] Fractional surface coverage, $f_{c,13}$ (%)	56	50	46	47	49
[j] $[A_0]_{23}$: RP-LC×IM (%)	66	64	70	90	73
[k] Fractional surface coverage, $f_{c,23}$ (%)	75	68	77	74	74
*[j] $[A_0]_{34}$: IM×MS (%)	26	30	38	19	28
*[k] Fractional surface coverage, $f_{c,34}$ (%)	15	13	39	12	20
[j] Degree of orthogonality, $[A_0]_{123}$ (%)	26	24	17	33	25
[k] Fractional surface coverage, $f_{c,123}$ (%)	29	16	11	17	18
[l] Practical peak capacity, $n'_{c,3D}$	61438	77203	93979	91128	80937
[m] Practical peak capacity, $n''_{c,3D}$	17731	12430	10169	15565	13974
[f] Separation power	15974	18529	15976	30072	20138

[a] Peak widths determined based on the number of 'slices' detected for each peak
[b] Calculated according to [46]
[c] Experimentally measured at half height and converted to 4σ peak width
[d] Corrected for ^{1}D undersampling
[e] Corrected for ^{1}D undersampling [49] and orthogonality [51]
[f] Indicative separation power, corrected for undersampling [49] and orthogonality [50]

[g] Calculated according to [49]
[h] Corrected for ^{2}D undersampling according to [49]
[i] Calculated according to [47,49]
[j] Calculated according to [50]
[k] Calculated according to [51]
* not used to calculate degree of orthogonality
[l] $n'_{c,3D} = {}^{1}n'_{c} \times {}^{2}n'_{c} \times {}^{3}n_{c}$
[m] Corrected for undersampling [49] and orthogonality [51]

Chapter 7

Conclusions and future recommendations

7.1 Conclusions

Due to the molecular complexity of plant (poly)phenolics, conventional analytical techniques frequently provide convoluted or inconclusive data for their mixtures, which complicates data interpretation. This is particularly true for hydrolysable tannins, which often occur as highly complex mixtures, the analytical characterisation of which is important to study their industrial and biological properties. It was, therefore, the goal of the work reported in this dissertation to explore the combination of several analytical techniques to overcome the challenges associated with analysing complex hydrolysable tannin mixtures. In particular, the hyphenation of one-dimensional and multidimensional liquid chromatography chromatographic separations with travelling wave ion mobility spectrometry (TWIMS) and high resolution mass spectrometry (HR-MS) detection were explored. Using various combinations of these methods, the previously unknown molecular complexity of hydrolysable tannins in tara and chestnut samples was revealed.

In the first part of the work, complementary separation methods in the form of hydrophilic interaction chromatography (HILIC) and reversed phase liquid chromatography (RP-LC) were used in combination with diode array and TWIMS-HR-MS detection for the detailed analysis of ellagitannins and gallotannins in chestnut and tara. This LC-UV-TWIMS-HR-MS$^{(E)}$ workflow provided comprehensive analytical information for tentative compound identification, thereby partially overcoming the limitation of a lack of commercial standards for hydrolysable tannins. Collisional cross sections (CCS's) obtained from the measured TWIMS arrival times for hydrolysable tannins are reported here for the first time, providing an additional criterion for compound identification. The accuracy of the CCS values was validated by comparative measurements obtained on a drift tube ion mobility spectrometry (DTIMS) instrument. Furthermore, identical arrival times and CCS values obtained between HILIC and RP-LC point to the high reliability of these measurements and, therefore, their suitability for identification purposes. The complimentary information provided by relative retention in two separation modes, as well as UV spectra, TWIMS arrival times and low and high collision energy high resolution MS data made it possible to tentatively characterise a large number of new molecular species and isomers in chestnut and tara.

In general, chestnut is characterised by a far greater diversity of hydrolysable tannins compared to tara. Chestnut contains ellagitannins, gallotannins, mixed ellagitannins species as well as hydrolysis products, whereas tara contained mainly gallotannins and a few low abundance species resulting from hydrolysis. A total of 38 molecular species comprising 136 isomers were tentatively identified in chestnut, of which 20 (comprising 78 isomers) are new species tentatively identified in the present work for the first time. In tara, 43 isomeric gallotannins were identified using the same workflow.

Good chromatographic performance was achieved by RP-LC for both chestnut and tara hydrolysable tannins. On the other hand, while HILIC showed similar good performance in addition to complementary selectivity for chestnut tannins, the performance of this mode was poor for the gallotannins present in tara. Typically, more hydrolysable tannin isomers were resolved using RP-LC, similar to what has been reported for condensed tannins. However, for chestnut, the choice of the separation mode depends on the species of interest. For example, the total number of gallotannin isomers detected is higher than the number of isomers resolved by either RP-LC or HILIC individually, revealing important selectivity differences between these techniques. The use of both one-dimensional methods, therefore, proved essential for the detailed analysis of chestnut hydrolysable tannins, and provided the motivation for the further development of comprehensive two-dimensional HILIC×RP-LC methods for hydrolysable tannin analysis.

Incorporating TWIMS into the LC-UV-MS workflow proved to be invaluable in terms of improving MS data quality and providing additional analytical information. The most important benefits include the option to obtain arrival time filtered mass spectra, which greatly facilitated tentative compound identification, as well as an improved duty cycle and, therefore, higher MS sensitivity. Ion mobility (IM) separation of hydrolysable tannins revealed ionic species with multiple arrival times, which are hypothesised to correspond to either prototropic isomers or conformers. Species displaying this behavior include the well-known dimeric vescalagin and castalagin species roburin A-D in chestnut, as well as the depsidic digalloyl quinic acid species identified in tara. The assignment of individual isomeric species between RP-LC and HILIC was also made possible based on IM data due to the highly reproducible nature of measured arrival times. Another prominent feature as revealed by LC-IM-MS is the formation of

multiply charged clusters of the ellagitannins vescalagin and castalagin under negative electrospray ionisation conditions. Finally, the diagnostic UV spectra of gallotannins comprising depsidic bond/s in combination with IM provided a simple approach to differentiate between structural isomers of this compound class.

Informed by the findings outlined above, on-line comprehensive HILIC×RP-LC separations were developed in order to exploit the high degree of orthogonality and resolving power (practical peak capacity) attainable for the separation of phenolic compounds using this approach. Further impetus for this work was provided by the observation that superior separation of specific classes of hydrolysable tannins were obtained by either RP-LC or HILIC. For this work, kinetic optimisation was used to derive optimal experimental conditions for the on-line coupling of HILIC and RP-LC separations. A methodology based on dilution of the first dimension (^{1}D) flow with an aqueous make-up provided informative group-type separations; practical peak capacities in excess of 1000 were measured for chestnut and tara separations. However, inherent constraints of on-line comprehensive LC×LC separations like on-column dilution and very fast second dimension (^{2}D) separations results in significantly fewer compounds detected in these samples compared to the optimised 1D RP-LC- and HILIC-UV-IM-MS methods. Nevertheless, this approach shows significant promise for high resolution screening of hydrolysable tannins in natural products, where the group-type separation obtained is especially useful.

In the final part of the work, the incorporation of TWIMS as a third separation step in combination with comprehensive on-line HILIC×RP-LC-HR-MS separation was explored as a potentially powerful analytical method for the improved separation of complex samples. Importantly, the fact that IM separation occurs on the millisecond timescale means that the technique meets the sampling criteria to be considered for a comprehensive 3-dimensional separation system. The application of the developed LC×LC×IM-MS method was demonstrated by the analysis of complex phenolic mixtures in wine, tea, chestnut and grape seed. Advantages of IM alluded to above, such as the improved MS duty cycle and arrival time filtered mass spectra also apply under these conditions. Furthermore, the additional separation step offered by IM enabled the distinction of isomeric procyanidin trimers based on unique arrival times, which could not be resolved using either LC×LC or HR-MS. Indeed, estimation of the

practical performance of the comprehensive HILIC×RP-LC×IM setup confirmed a thirteen-fold increase in separation performance compared to HILIC×RP-LC. Several instrument-related limitations were also noted in this work. In the first instance, on the instrument used here, the combination of IM separations to correspond to the MS scan time artificially results in the under-sampling of narrow ^{2}D RP-LC peaks. Furthermore, undersampling of ion mobility peaks due to the slow sampling rate of time-of-flight (TOF) analyser and diffusion of ions during IM separation further impacted negatively on the separation performance. Despite these instrument specific limitations, the LC×LC×IM-MS methodology reported here for the first time shows significant promise for the improved analysis of highly complex samples; although demonstrated in the present work for phenolic compounds, in principle the same approach can find beneficial application for any LC-amenable samples.

7.2 Future recommendations

For future research, the relatively large number of hydrolysable tannins tentatively identified for the first time should be isolated and confirmed using NMR. It should be emphasised that the new hydrolysable tannin species reported in this thesis were tentatively identified, albeit with a relatively high degree of certainty. For example, it has been shown that the combination of UV and ion mobility spectrometry provide some insight into the isomeric composition of especially tara gallotannins. However, for unambiguous assignment of the structures of these compounds, including their stereochemistry, NMR is required. From the chromatographic results obtained during this work, upscaling the existing RP-LC and HILIC analytical techniques would be an effective starting point in such research. Our findings also point to some of the most interesting compounds that may realistically be isolated for NMR characterisation.

It is evident that further investigation is needed into the reason(s) for specific hydrolysable tannin (and phenolic) species displaying multiple arrival times. To investigate this aspect, several approaches may be used, including detailed molecular modeling [1], the combination of IM with different spectroscopic techniques to deduce structural information [2], and molecular derivatisation to investigate the chemical features possibly responsible for this occurrence [3].

Further work in the area of LC×LC×IM should include evaluation of the performance of alternative IM instruments such as drift tube ion mobility spectrometry in such systems. Many further applications of this novel 3-dimensional analytical method can be envisaged.

Another area of research which is especially relevant for high-throughput laboratories, is the possibility of replacing LC-MS methods with IM-MS, which can drastically increase analytical throughput and reduce cost [4].

7.3 References

1. Zheng X, Renslow RS, Makola MM, Webb IK, Deng L, Thomas DG, Govind N, Ibrahim YM, Kabanda MM, Dubery IA, Heyman HM, Smith RD, Madala NE, Baker ES (2017) Structural Elucidation of cis/trans Dicaffeoylquinic Acid Photoisomerization Using Ion Mobility Spectrometry-Mass Spectrometry. J Phys Chemsitry Lett 8:1381–1388

2. Warnke S, Seo J, Boschmans J, Sobott F, Scrivens JH, Bleiholder C, Bowers MT, Gewinner S, Schöllkopf W, Pagel K, Von Helden G (2015) Protomers of benzocaine: Solvent and permittivity dependence. J Am Chem Soc 137:4236–4242. doi: 10.1021/jacs.5b01338

3. Kuhnert N, Yassin GH, Jaiswal R, Matei MF, Grün CH (2015) Differentiation of prototropic ions in regioisomeric caffeoyl quinic acids by electrospray ion mobility mass spectrometry. Rapid Commun Mass Spectrom 29:675–680

4. Hädener M, Kamrath MZ, Weinmann W, Groessl M (2018) High-Resolution Ion Mobility Spectrometry for Rapid Cannabis Potency Testing. Anal Chem 90:8764–8768. doi: 10.1021/acs.analchem.8b02180